







```

Qy 2041 ATGAAATTCATATCTGTGAACAGACAGAGTTCAAAAACTAAGCATCCAGAGCCCT 2100
Db 2041 ATGAAATTCATATCTGTGAACAGACAGAGTTCAAAAACTAAGCATCCAGAGCCCT 2100
Qy 2101 GAGCAGCTTTCAGCACTGAGAGTGAAGAGCGCTGTTTTTAAATAACAGACAGACAG 2160
Db 2101 GAGCAGCTTTCAGCACTGAGAGTGAAGAGCGCTGTTTTTAAATAACAGACAGACAG 2160
Qy 2161 TCAAGGTGTTTTCAGAGCCCTGAGAGGAGGAGCGAGGTCTCCGACAGGTCCTG 2220
Db 2161 TCAAGGTGTTTTCAGAGCCCTGAGAGGAGGAGCGAGGTCTCCGACAGGTCCTG 2220
Qy 2221 GGTGACTCTTCTGTGAGGCTTTTACCTCTGAGTGAACCTCCAGAGCCCGGGG 2280
Db 2221 GGTGACTCTTCTGTGAGGCTTTTACCTCTGAGTGAACCTCCAGAGCCCGGGG 2280
Qy 2281 CCGCAGCCCGCCCTCTGTGAGCGCTGAGGAGGCTGCTGTGAGTGAACAGACAG 2340
Db 2281 CCGCAGCCCGCCCTCTGTGAGCGCTGAGGAGGCTGCTGTGAGTGAACAGACAG 2340
Qy 2341 CGAAGCTTCTGTGATCATGCGCGCTCCCGAGAGGAGGAGCTTTGCTTTGTAC 2400
Db 2341 CGAAGCTTCTGTGATCATGCGCGCTCCCGAGAGGAGGAGCTTTGCTTTGTAC 2400
Qy 2401 ATTTTCGAAACTACAGTTAAAGCAGAGAGTCTGTTTTCAGAAAAGTTTCAAGGAGAG 2460
Db 2401 ATTTTCGAAACTACAGTTAAAGCAGAGAGTCTGTTTTCAGAAAAGTTTCAAGGAGAG 2460
Qy 2461 GGCAGATTATCAAAAACATTTGTTTCAGAGAGAGGAGCATAGTTTACAGCCTACAG 2520
Db 2461 GGCAGATTATCAAAAACATTTGTTTTCAGAGAGAGGAGCATAGTTTACAGCCTACAG 2520
Qy 2521 CGTACACATATCCGCTGCTGAGAAACACAGCATTTTATCTATTTTATTTTAA 2580
Db 2521 CGTACACATATCCGCTGCTGAGAAACACAGCATTTTATCTATTTTATTTTAA 2580
Qy 2581 GGTGGTCTTATCTTCTAATTAAGATTAAATGTCAAACTGTGACAAATATATA 2640
Db 2581 GGTGGTCTTATCTTCTAATTAAGATTAAATGTCAAACTGTGACAAATATATA 2640
Qy 2641 ATTTAATTTTCAAAATGACAAAAAATTTTTTTTTT 2679
Db 2641 ATTTAATTTTCAAAATGACAAAAAATTTTTTTTTT 2679
RESULT 2
AX877194 2639 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 12099 from Patent EP1074617.
DEFINITION AX877194
ACCESSION AX877194
VERSION AX877194.1 GI:40031930
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Oda,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primers for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 12099 07-FEB-2001;
Research Association for Biotechnology (JP)
LOCATION/Qualifiers
1..2639
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
79..2037
/notes="unassigned protein product"
/codon_start=1
/protein_id="CAE89614.1"
/db_xref="GI:40031931"
/translation="MERPEBKGSPPPQWGRLLRLGAEBSGPHVLLRKREWTIGRR

```

## ORIGIN

```

Query Match 95.8%; Score 2566.4; DB 6; Length 2639;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2613; Conservative 0; Mismatches 1; Indels 36; Gaps 1;

Qy 17 GCCGAAATGCTCTTGAACAGCGCGCGCGCGCGGTTCCGGTTCGGCGCGCGCG 76
Db 5 GCCGAAATGCTCTTGAACAGCGCGCGCGCGCGCGGTTCCGGTTCGGCGCGCGCG 64
Qy 77 GGATGTAATCCCGATGAGCGCGCGCGAGAGGAGAGAGTGCCTCCGCGCGCGCCT 136
Db 65 GGATGTAATCCCGATGAGCGCGCGCGAGAGAGAGAGTGCCTCCGCGCGCGCCT 124
Qy 137 GGGGACGGCTCTGCTGCGCGCGCGAGAGGAGAGAGCGCGCGCTCTCTGAGAGAG 196
Db 125 GGGGACGGCTCTGCTGCGCGCGCGAGAGAGAGAGCGCGCGCTCTCTGAGAGAG 184
Qy 197 GGGAGTGAACATCGCGCGAGAGAGAGTTCGACCTTCTTCCCGACATTAACCTG 256
Db 185 GGGAGTGAACATCGCGCGAGAGAGAGTTCGACCTTCTTCCCGACATTAACCTG 244
Qy 257 TCTTGAAGATCACTGTAAGATTGTAAGTGAATAAATCAGGTGAGTGAACCTGAG 316
Db 245 TCTTGAAGATCACTGTAAGATTGTAAGTGAATAAATCAGGTGAGTGAACCTGAG 304
Qy 317 ATACGAGACAGAGTGAACAGTATTAACAGTGAAGTCTTGAAGAGAGAGATGCG 376
Db 305 ATACGAGACAGAGTGAACAGTATTAACAGTGAAGTCTTGAAGAGAGAGATGCG 364
Qy 377 CTTTACAGAGTGGAGTGCATCTTCTGTTAGAGAGAGATGAACCGAGAACAG 436
Db 365 CTTTACAGAGTGGAGTGCATCTTCTGTTAGAGAGAGATGAACCGAGAACAG 424
Qy 437 TGGCATACCTCTATGAATCTTTAAGTGAAGAGAGAGTGAACAGAGATCTTTGAG 496
Db 425 TGGCATACCTCTATGAATCTTTAAGTGAAGAGAGAGTGAACAGAGATCTTTGAG 480
Qy 497 CTAAAGAGAAATGTGTCATGGAGCAAAAGATACCTCAGGTGAGTGAAGGCGAG 556
Db 481 -----GATACCTCAGGTGAGTGAAGGCGAG 508
Qy 557 GGGCGCATCCCGGGTCCCTCCGCTGTCGCGCGCGACATCAGGTGCTTTGAGAACAC 616
Db 509 GGGCGCATCCCGGGTCCCTCCGCTGTCGCGCGCGACATCAGGTGCTTTGAGAACAC 568
Qy 617 AGCCATCAACATGACGTCAGACTCTTCCCGACAGCTCGGCTCTTCCAGAGAGCTT 676
Db 569 AGCCATCAACATGACGTCAGACTCTTCCCGACAGCTCGGCTCTTCCAGAGAGCTT 628
Qy 677 CTCCTGAGAGGAGAGAGGCTTCTCAAGTTGAGGTCTGAGGAGTGAAGTCTCCCTTA 736
Db 629 CTCCTGAGAGGAGAGAGGCTTCTCAAGTTGAGGTCTGAGGAGTGAAGTCTCCCTTA 688
Qy 737 AAGAAAGTGTCCCTCTGAGAGTGAAGTCTCAGCTTTGCTCAGGTCTCCAG 796
Db 689 AAGAAAGTGTCCCTCTGAGAGTGAAGTCTCAGCTTTGCTCAGGTCTCCAG 748
Qy 797 AAGAAAGTGTCCCTCTTGTGTTGAACCCAGAGATCAGAGATTTGAGAGCCG 856
Db 749 AAGAAAGTGTCCCTCTTGTGTTGAACCCAGAGATCAGAGATTTGAGAGCCG 808

```





QY 1877 GAGACACCGTTCTGTGTACTGTCTGTGCTGTGCGCCTTCCGTGACCTGACCTATCACT 1936  
 DB 1829 GAAACACCGCTTCTGTGTACTGTCTGTGCGCTGTGCGAGCTTCCGTGAGCTGACCTATCACT 1888  
 QY 1937 ATCGGAGAACATTTCTGTGCTTCCGAGTTGCGAGTGGCCGTAACATCCGCTCCGAGCTGCT 1996  
 DB 1889 ATCGGAGAACATTTCTGTGCTTCCGAGTTGCGAGTGGCCGTAACATCCGCTCCGAGCTGCT 1948  
 QY 1997 ACTGGGGCGGTAACTGCGCAGCTCAGGTGAAAGCTCAGCAGCGCATGAATTCATCATTA 2056  
 DB 1949 ACTGGGGCGGTAACTGCGCAGCTCAGGTGAAAGCTCAGCAGCGCATGAATTCATCATTA 2008  
 QY 2057 TCTGTGAACAAGCAAGCTCAAAATACTAGCATCCGAGCGCTGAGCGAGCTTTCAGCAGC 2116  
 DB 2009 TCTGTGAACAAGCAAGCTCAAAATACTAGCATCCGAGCGCTGAGCGAGCTTTCAGCAGC 2068  
 QY 2117 TGGAGGTGAAGAGAGCGTGTATTTAAATATAGAGACAGCAAGTAAAGTGTATTCACA 2176  
 DB 2069 TGGAGGTGAAGAGAGCGTGTATTTAAATATAGAGACAGCAAGTAAAGTGTATTCACA 2128  
 QY 2177 GCCCCTTGAAGGAGAGGAGCGAGGCTCTCCGACAGGTCTCTGGAGGTGAATCTTCTGTGG 2236  
 DB 2129 GCCCCTTGAAGGAGAGGAGCGAGGCTCTCCGACAGGTCTCTGGAGGTGAATCTTCTGTGG 2188  
 QY 2237 AGCTTTTTCCTCTGAGTGAACCTTCCCGAGAGCCCGGGGGCGCGAGCCCGCTTC 2296  
 DB 2189 AGCTTTTTCCTCTGAGTGAACCTTCCCGAGAGCCCGGGGGCGCGAGCCCGCTTC 2248  
 QY 2297 TGGTGAAGCGCTGGGCGAGGCTGCTGTGTGATCAGACAGAGACGAAAGCTTTCTGTAA 2356  
 DB 2249 TGGTGAAGCGCTGGGCGAGGCTGCTGTGTGATCAGACAGAGACGAAAGCTTTCTGTAA 2308  
 QY 2357 CATGCGCCCGCTCCCGCGAGAGGCGAGTGTGTCTTTTGTACATTTTCCGAAACTACA 2416  
 DB 2309 CATGCGCCCGCTCCCGCGAGAGGCGAGTGTGTCTTTTGTACATTTTCCGAAACTACA 2368  
 QY 2417 GTTAAAGCAGAGTCTGTTCAGGAAAGTTCAGGAGGAGGAGGAAAGTTCATCAAAA 2476  
 DB 2369 GTTAAAGCAGAGTCTGTTCAGGAAAGTTCAGGAGGAGGAGGAAAGTTCATCAAAA 2428  
 QY 2477 ACATTTTTCAGAGAGAGGAGCATTAAGTTTACAGCTTACAGAGATACACAAATCTCG 2536  
 DB 2429 ACATTTTTCAGAGAGAGGAGCATTAAGTTTACAGCTTACAGAGATACACAAATCTCG 2488  
 QY 2537 CTGCTGGGAGAAACACAGATTTTATCTATTTTATTTTATTTAGTGTCTATCT 2596  
 DB 2489 CTGCTGGGAGAAACACAGATTTTATCTATTTTATTTTATTTAGTGTCTATCT 2548  
 QY 2597 TCTAATTAAGATTTAATGTCACAAACTGTAGACAAATTAATTAATTAATTAACAAA 2656  
 DB 2549 TCTAATTAAGATTTAATGTCACAAACTGTAGACAAATTAATTAATTAATTAACAAA 2608  
 QY 2657 TTGACAAAAA 2666  
 DB 2609 TTGACAAAAA 2618  
 RESULT 4  
 AK027687  
 LOCUS Homo sapiens cDNA FLJ14781 fib, clone NT2RP4000455, weakly similar  
 DEFINITION to TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0.  
 ACCESSION AK027687.1 GI:14042552  
 VERSION AK027687  
 KEYWORDS Oligo cloning; fib (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,  
 Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,  
 Sekine, M., Oba, Y., Nishi, T., Shibahara, T., Tanaka, T.,

Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,  
 Nagahara, K., Murakami, K., Yasuda, T., Iwaguchi, T., Nagatsuma, M.,  
 Shiratori, A., Sudo, H., Hosokawa, T., Kaku, Y., Kodaira, H., Kondo, H.,  
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,  
 Kikawa, E., Omura, Y., Abe, K., Kanthara, K., Katsube, N., Sato, K.,  
 Tanikawa, M., Yamazaki, M., Nishimura, K., Ishibashi, T., Yamashita, H.,  
 Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,  
 Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,  
 Yoshida, M., Hoshino, T., Kusano, J., Kanehori, K., Takahashi, F.,  
 Hara, H., Tanase, T., Nomura, Y., Togaya, S., Komai, F., Hara, R.,  
 Tanouchi, K., Arima, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A.,  
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,  
 Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Sato, N., Takami, S.,  
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mazonuchi, H.,  
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,  
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,  
 Kunagai, A., Itakura, S., Fukunumi, Y., Fujimori, Y., Komiyama, M.,  
 Tashiro, H., Tanigami, A., Fujitani, T., Oho, T., Yamada, K., Fujii, Y.,  
 Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N.,  
 Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,  
 Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunawa, K.,  
 Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,  
 Oyama, M., Hara, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,  
 Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,  
 Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R.,  
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.  
 Complete sequencing and characterization of 21,243 full-length  
 human cDNAs  
 Nat. Genet. 36 (1), 40-45 (2004)  
 14702039  
 2  
 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
 Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,  
 Matsunawa, A., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,  
 Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuno, Y. and Sasaki, N.  
 NEDO human cDNA sequencing project  
 Unpublished  
 3 (bases 1 to 2639)  
 Isogai, T. and Otsuki, T.  
 Direct Submission  
 Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology; cDNA library construction,  
 5' - & 3' - end one pass sequencing and clone selection; Helix  
 Research Institute (supported by Japan Key Technology Center etc.)  
 and Department of Virology, Institute of Medical Science,  
 University of Tokyo.  
 Location/Qualifiers  
 1..2639  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2RP4000455"  
 /cell\_line="NT2"  
 /cell\_type="teratocarcinoma"  
 /note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal  
 precursor cells after 2-weeks retinoic acid (RA)  
 induction."  
 79..2037  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB55297.1"  
 /db\_xref="GI:14042553"  
 /translation="MRRPREGKSGSPPPQWGRLLRLGABEGEPHYLLRKREWTGRR  
 GGLSPSPNKLVSQDHCRIVVDEKSDQVLEEDTSGTIVNKLKVKQCTPQGTGV  
 IYLVKRNPEHNVAALYSELSEKQMTDESFPDTSAGAGAGADPRVPPSPAYQVCF  
 EBPQPSITSDLPPTASGASTBPSAPGKRESSSCGSGCGGCTISPKSGSPVASEVSSF  
 ASALPDKRTSFSLEPDQEDLEPKKRGDGDLDINQLLVAQPRNAQVYHEV  
 RAAAGPDKMEFTLTCIQDILHDCVSLQPCNHTCAACYSQMWESSLSLCPTRCRGV











Db 1825 AGAATCTTCCTGCTCCGAGTTCAGAGGCGGTAACATCCGCTCTGACTGCTGAG 1884  
 Qy 2003 GCCGTAATGCGCGCACTCAGGTGAAAGCTCACCACGCGCATGAAATTCATATCTG 2062  
 Db 1885 GCGGTACGCGCGCACTCAGGTGAAAGCTCACCACGCGCATGAAATTCATATCTG 1944  
 Qy 2063 AACAGCAAGGTTCAAAAATCAAGCATCCAGAGGCGCTTGAAGAGCTTTCAGCATGAG 2122  
 Db 1945 AACAGCAAGGTTCAAAAATCAAGCATCCAGAGGCGCTTGAAGAGCTTTCAGCATGAG 2004  
 Qy 2123 TGAAGAGAGGTTGTTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2182  
 Db 2005 TGAAGAGAGGTTGTTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2064  
 Qy 2183 TGAAG 2242  
 Db 2065 TGAAG 2124  
 Qy 2243 TTAACCTCTGATGAG 2302  
 Db 2125 TTAACCTCTGATGAG 2184  
 Qy 2303 GCGCTGAG 2362  
 Db 2185 GCGCTGAG 2244  
 Qy 2363 GCGCTGAG 2422  
 Db 2245 GCGCTGAG 2304  
 Qy 2423 GAG 2482  
 Db 2305 GAG 2364  
 Qy 2483 TTTCAAG 2542  
 Db 2365 TTTCAAG 2424  
 Qy 2543 GGAAG 2602  
 Db 2425 GGAAG 2484  
 Qy 2603 AAGATTTAAATGTCACAACTGATGACAAATTAATTAATTAATTAATTAATTAATTA 2662  
 Db 2485 AAGATTTAAATGTCACAACTGATGACAAATTAATTAATTAATTAATTAATTAATTA 2544  
 Qy 2663 AAAA 2666  
 Db 2545 AAAA 2548  
 RESULT 6  
 LOCUS AX877222 3138 bp DNA linear PAT 17-DEC-2003  
 DEFINITION Sequence 12127 from Patent EP1074617.  
 ACCESSION AX877222  
 VERSION AX877222.1 GI:40031958  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 Ota, T., Iwaga, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.  
 Primers for synthesizing full-length cDNA and their use  
 Patent: EP 1074617-A 12127 07-FEB-2001  
 JOURNAL Research Association for Biotechnology (JP)  
 FEATURES  
 Location/Qualifiers  
 1..3138  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

CDS  
 65..1936  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CA89626.1"  
 /db\_xref="GI:40031958"  
 /translation="MERPEBKGSPPPQPMGRLLRLGAESEPHVLRKKEWTGRR  
 GCDLSPSPNKLVSADHCRIVVDEKSGVLTEDTSTGIVNKLVRKQCPLOTGDV  
 IYLYVRKNEPHEHNYALYIESLSEKQMTQESFEMVPCCVQAGLKLGSBDPTLASQ  
 SIVITGSGGGGSI SPKSGSPVYASDEVSFASLPDRKAFSPSLPQDEDLPEBEKK  
 MRGDDDLINGOLIVAOPRNAOTVHEDVAAAKPDMEEETLCIIICODLHDVSL  
 QPCNHTCAACYSQWMEERSLCPKPCRPVRIKCNHLLNNLYEALVLIQDPKSPSEED  
 VQSDARAKLTQDMLQKVRKRSFDESGSSBDLELSDVBSSESDIQPYVQCCE  
 YRDAQPPHCPAPESPGAPQALGDAPPTVSILTVIVQDVCPLOQSHALCTCFDP  
 MPDRARQEDPRAVAPQCAVCLQPFGLHWKCTRCYGLAPFCELNADKLDGV  
 LNNYSYSDILKNTLATRGLTKMMLTESLVALQRYVFLSDYVETDPTVLYCGLR  
 SFRELTVYOQOINI PASLPLVAVTSRPPCYMGRNCRTOVKAHAMKFNHICQTRFKN"

ORIGIN  
 Query Match 84.4%; Score 2262.4; DB 6; Length 3138;  
 Best Local Similarity 93.0%; Pred. No. 0;  
 Matches 2454; Conservative 0; Mismatches 61; Indels 125; Gaps 3;  
 27 CTCTTGA CAGGCGGCGGCGGCGGCGGCTTCGGGTTCCGGCGGCGGCGGCGGAGTGTGAT 86  
 1 CTCTTGA CAGGCGGCGGCGGCGGCGGCTTCGGGTTCCGGCGGCGGCGGCGGAGTGTGAT 60  
 Qy 87 CCCGATGAGAGCGGCCCGGAG 146  
 Db 61 CCCGATGAGAGCGGCCCGGAG 120  
 Qy 147 CCTGCTGTCGGGCGCGGAG 206  
 Db 121 CCTGCTGTCGGGCGCGGAG 180  
 Qy 207 CATGCGGCGGAG 266  
 Db 181 CATGCGGCGGAG 240  
 Qy 267 TCACGTGAGAGATTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326  
 Db 241 TCACGTGAGAGATTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 Qy 327 CAGTGAAG 386  
 Db 301 CAGTGAAG 360  
 Qy 387 TGGGAGATGATCACTACTTGTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446  
 Db 361 TGGGAGATGATCACTACTTGTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 Qy 447 CTATGAATCTTTAAGTGAAG 506  
 Db 421 CTATGAATCTTTAAGTGAAG 480  
 Qy 507 AAATGTTTCCATGAG 566  
 Db 481 CTGTTGTCCTCAGGCTGCTTAAATCTCTGAGATCAAGT----- 520  
 Qy 567 CCGGATCCCTCCGTCGTGCGCCGCGCACTCAGAGTGTGCTTTGAGAGAGAGAGAGAGAGAG 626  
 Db 521 ---GATCTCCACCTTGGGCTCCCAAGATATGATTTCA----- 559  
 Qy 627 ATGACGTCAAGACTTCTCCACAGGCTGAGCTCTTCCACGAGAGGCTTCTCTGACAG 686  
 Db 560 ----- 559  
 Qy 687 GCGAGAGCGTTCTCCAGTGTGGGTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 746  
 Db 560 -----GGGTTGAG 597  
 Qy 747 TCCCTCTGTGAG 806  
 Db 598 TCCCTCTGTGAG 657

QY 807 TGGCTCTTTTGGTGGAAACCCAGATCAGAGATTGGAGCCGTGAAGAA 866  
 DB 658 TGGCTCTTTTGGTGGAAACCCAGATCAGAGATTGGAGCCGTGAAGAA 717  
 QY 867 AATGAGAGAGATGAGGAGCTTGAACCTGAACGGGCAATTGTTGGTGCACAAACGGCTAG 926  
 DB 718 AATGAGAGAGATGAGGAGCTTGAACCTGAACGGGCAATTGTTGGTGCACAAACGGCTAG 777  
 QY 927 AATGCTCCAAACCTGTCACAGAGAGCTCAGAGAGCGGCTGGAGACCAAGATGGA 986  
 DB 778 AATGCTCCAAACCTGTCACAGAGAGCTCAGAGAGCGGCTGGAGACCAAGATGGA 837  
 QY 987 GAGAGAGCTGACATGATCATCTGACAGAGCTGCTGACAGATGGGTGATTGAGGC 1046  
 DB 838 GAGAGAGCTGACATGATCATCTGACAGAGCTGCTGACAGATGGGTGATTGAGGC 897  
 QY 1047 CTGCATGACACAGTTCTGCGCGCTTGTACTGCGGCTGGATGAGAGCTGCTGCTG 1106  
 DB 898 CTGCATGACACAGTTCTGCGCGCTTGTACTGCGGCTGGATGAGAGCTGCTGCTG 957  
 QY 1107 TCTTACCTGCGCTGTCCTGCGGAGCGGATCTGTAAACCAATCTCTCAACCTCGT 1166  
 DB 958 TCTTACCTGCGCTGTCCTGCGGAGCGGATCTGTAAACCAATCTCTCAACCTCGT 1017  
 QY 1167 GGAAGCATCTCATGACAGATCAGACAAAGTCCGAGTGAAGAGATGGAAGTAT 1226  
 DB 1018 GGAAGCATCTCATGACAGATCAGACAAAGTCCGAGTGAAGAGATGGAAGTAT 1077  
 QY 1227 GATGCCAGGAAATAAATCACTCAAGACATGCTGACAGCCCAAGTGGCGCTTTTTC 1286  
 DB 1078 GATGCCAGGAAATAAATCACTCAAGACATGCTGACAGCCCAAGTGGCGCTTTTTC 1137  
 QY 1287 TGATGAAGAGAGATTCAGAGAGCTGCTGAGCTGTCAAGCTTGAACATGATGCTTC 1346  
 DB 1138 TGATGAAGAGAGATTCAGAGAGCTGCTGAGCTGTCAAGCTTGAACATGATGCTTC 1197  
 QY 1347 AGACATTAAGCCAGCCTAGCTGCTGAGCGGAGTCCGAGTGAAGAGAGAGCGGC 1406  
 DB 1198 AGACATTAAGCCAGCCTAGCTGCTGAGCGGAGTCCGAGTGAAGAGAGAGCGGC 1257  
 QY 1407 GAGCCTCTCCCACTGCGCAGCAACCGAGGCGAGAGAGCCCAAGGCTTGGGGGA 1466  
 DB 1258 GAGCCTCTCCCACTGCGCAGCAACCGAGGCGAGAGAGCCCAAGGCTTGGGGGA 1317  
 QY 1467 TGACCTCTCAAGCTGCTGACCTGACAGAGAGTCCAGGATTAAGTCTGCTTGA 1526  
 DB 1318 TGACCTCTCAAGCTGCTGACCTGACAGAGAGTCCAGGATTAAGTCTGCTTGA 1377  
 QY 1527 AGGAGAGCAGCGCTGCTGACCTGCTTCCAGGCCATGCCAGACCGGAGAGCGG 1586  
 DB 1378 AGGAGAGCAGCGCTGCTGACCTGCTTCCAGGCCATGCCAGACCGGAGAGCGG 1437  
 QY 1587 CGAGCAGAGACCGCGCTGCGCCCTCAGCAGTGGCGTCTGCTGAGCTTTCTGCCA 1646  
 DB 1438 CGAGCAGAGACCGCGCTGCGCCCTCAGCAGTGGCGTCTGCTGAGCTTTCTGCCA 1497  
 QY 1647 CTTGTAAGTGGGCTGACCGCGAGACCGGCTGCTAAGCTGCTGAGCCCTTTTGTAGCT 1706  
 DB 1498 CTTGTAAGTGGGCTGACCGCGAGACCGGCTGCTAAGCTGCTGAGCCCTTTTGTAGCT 1557  
 QY 1707 CAACCTGGGTGACAAATGCTGACAGCGGCTGTGAACCAACAGCTAAGATCAACAT 1766  
 DB 1558 CAACCTGGGTGACAAATGCTGACAGCGGCTGTGAACCAACAGCTAAGATCAACAT 1617  
 QY 1767 CTTGAAGATTAAGTGGCAACAGAGGTTTGAATGAGAAAAATCATGTTGACGAGAGCT 1826  
 DB 1618 CTTGAAGATTAAGTGGCAACAGAGGTTTGAATGAGAAAAATCATGTTGACGAGAGCT 1677  
 QY 1827 CTTGAGCTTTCAGCGGAGAGTGTCTGCTGTGATTAAGAGTCAAGGAGACACCGT 1886  
 DB 1678 CTTGAGCTTTCAGCGGAGAGTGTCTGCTGTGATTAAGAGTCAAGGAGACACCGT 1737

QY 1887 TCTGTTTACCTGCTGAGCTGCGCAGCTTCCGTGAGCTGACCTATCAGATCGAGAA 1946  
 DB 1738 TCTGTTTACCTGCTGAGCTGCGCAGCTTCCGTGAGCTGACCTATCAGATCGAGAA 1797  
 QY 1947 CATTCCTGCTTCCAGTTCACAGTGGCGGTAAATCCGCTCTGACTGCTGAGGCGG 2006  
 DB 1798 CATTCCTGCTTCCAGTTCACAGTGGCGGTAAATCCGCTCTGACTGCTGAGGCGG 1857  
 QY 2007 TAACTCCGCACTCAGTGTAAAGCTCAACGCCATGAATTAATCATATCTGAGAA 2066  
 DB 1858 TAACTCCGCACTCAGTGTAAAGCTCAACGCCATGAATTAATCATATCTGAGAA 1917  
 QY 2067 GACAAGTTCAAAAATAAGCATCCAGAGGCTTGAAGCTTTCAGACTGAGAGTGA 2126  
 DB 1918 GACAAGTTCAAAAATAAGCATCCAGAGGCTTGAAGCTTTCAGACTGAGAGTGA 1977  
 QY 2127 GAGAGGTTTAAATTAAGACAGACAGTCAAGTGTTCACAGCCCTCTGAG 2186  
 DB 1978 GAGAGGTTTAAATTAAGACAGACAGTCAAGTGTTCACAGCCCTCTGAG 2037  
 QY 2187 GGAAGGAGCAGAGTCTCCGACAGTGTCTGAGGTGACTTCTGTGAGAGCTTTTAC 2246  
 DB 2038 GGAAGGAGCAGAGTCTCCGACAGTGTCTGAGGTGACTTCTGTGAGAGC--TTTAC 2095  
 QY 2247 CTTGAGTGAAGACCTCCCAAGGCGCGGAGCGCGAGCCCTCTGAGTGAAGC 2306  
 DB 2096 CTTGAGTGAAGACCTCCCAAGGCGCGGAGCGCGAGCCCTCTGAGTGAAGC 2155  
 QY 2307 TGGGCAAGGCTGCTGAGTGCATCAGACAGACAGACAGCTTCTGTAACATGCGCG 2366  
 DB 2156 TGGGCAAGGCTGCTGAGTGCATCAGACAGACAGACAGCTTCTGTAACATGCGCG 2215  
 QY 2367 TCCGCGGAGAGAGGAGTGTGCTCTTTTGTATCTTCCGAACTAAGTTAAAGCAG 2426  
 DB 2216 TCCGCGGAGAGAGGAGTGTGCTCTTTTGTATCTTCCGAACTAAGTTAAAGCAG 2275  
 QY 2427 AAGTCTGTTTCAAGAAAGTTTCAAGGAGAGAGGCAAGTTTATCAAAAATGTTTC 2486  
 DB 2276 AAGTCTGTTTCAAGAAAGTTTCAAGGAGAGAGGCAAGTTTATCAAAAATGTTTC 2335  
 QY 2487 AGGAGAGGAGCATTAAGTTTACAGCTTACAGAGCTGACAGATATCTGCTGAGGAA 2546  
 DB 2336 AGGAGAGGAGCATTAAGTTTACAGCTTACAGAGCTGACAGATATCTGCTGAGGAA 2395  
 QY 2547 AACACAGCATTTTATCTATTTTATTTTATTAAGTTTGTGCTTATCTTAAATGA 2606  
 DB 2396 AACACAGCATTTTATCTATTTTATTTTATTAAGTTTGTGCTTATCTTAAATGA 2455  
 QY 2607 TTTAAATGCAAACTGTAGCAAAATTAATTAATTAATTAATTAATTAATGA 2666  
 DB 2456 TTTAAATGCAAACTGTAGCAAAATTAATTAATTAATTAATTAATTAATGA 2515

RESULT 7  
 BD156548  
 LOCUS  
 DEFINITION  
 Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION  
 BD156548.1 GI:27862306  
 VERSION  
 JP 2002191363-A/11391.  
 KEYWORDS  
 Homo sapiens (human)  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 3138)  
 Ota,T., Iseoka,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Iehli,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 PN JP 2002191363-A/11391  
 PD 09-JUL-2002



Qy	1887	CTGTGTTACTACGCTCTGTGGCTTGGCCGAGGCTCCGCTAGCTGACCTTACGTTATCCGCAAGA	19446
Dp	1738	TCGTGTTACTACGCTCTGTGGCTTGGCCGAGGCTCCGCTAGCTGACCTTACGTTATCCGCAAGA	17979
Qy	1947	CATTCTCTGCTTCCGAGTTGCCAGTGGCCGTAACATCCGCTCTGACTGACTCTACTGGGCGC	20060
Dp	1798	CATTCTCTGCTTCCGAGTTGCCAGTGGCCGTAACATCCGCTCTGACTGACTCTACTGGGCGC	1857
Qy	2007	TAACTGCCCCGCACTCAGGTGAAAGCTTCACACGCCATGAAATTCAATCATATCTGTGAACA	20666
Dp	1858	TAACTGCCCCGCACTCAGGTGAAAGCTTCACACGCCATGAAATTCAATCATATCTGTGAACA	19174
Qy	2067	GACAAAGTTCAAAAACCTAAGCACTCCAGAGGCCCTGAGAGGCTTTCAAGCACTGAGAGGTAA	21260
Dp	1918	GACAAAGTTCAAAAACCTAAGCACTCCAGAGGCCCTGAGAGGCTTTCAAGCACTGAGAGGTAA	19777
Qy	2127	GAGAGCGTGTTTTTTAAATATACAGAGACAAGCACGTCAAGGTGTATTCAACAGCCCTCTAG	21868
Dp	1978	GAGAGCGTGTTTTTTAAATATACAGAGACAAGCACGTCAAGGTGTATTCAACAGCCCTCTAG	20376
Qy	2187	GGAAGGGAACGCAAGGCTCTCCACACAGTGTCTGTGGGTGATCTTTCTGTGAGACTTTTAC	22484
Dp	2038	GGAAGGGAACGCAAGGCTCTCTCCACAGTGTCTGTGGGTGATCTTTCTGTGAGAC--TTTAC	20959
Qy	2247	CCTCTGAGTGAAGACCTCTCCCGCAGAGCCCGGGGGGCGGAGGCCCGCCCTCTGTGTAGAGGC	23060
Dp	2096	CCTCTGAGTGAAGACCTCTCCCGCAGAGCCCGGGGGGCGGAGGCCCGCCCTCTGTGTAGAGGC	21555
Qy	2307	TGGGCAAGGACTCGTGTGTGCATCAGCAGACAGAGCGAAAGCCTTCTGTAAATATGGGCGC	23666
Dp	2156	TGGGCAAGGACTCGTGTGTGCATCAGCAGACAGAGCGAAAGCCTTCTGTAAATATGGGCGC	2215
Qy	2367	TCCGCGCCAGAGGGGCAAGTTTGTCTTTTGTACATTTTCCGAAACTACAGTTAAAGAG	24280
Dp	2216	TCCGCGCCAGAGGGGCAAGTTTGTCTTTTGTACATTTTCCGAAACTACAGTTAAAGGCG	22757
Qy	2427	AAGTCTGTATTCAAGGAAAGTTTCAAGGGAAGAGGGCAAGTTTATCAAAAATTGTCTC	24860
Dp	2276	AAGTCTGTATTCAAGGAAAGTTTCAAGGGAAGAGGGCAAGTTTATCAAAAATTGTCTC	23353
Qy	2487	AGGAGAAAGGAGCATTAAGTTTACAGCCTTACAGAGCGTACACAATATCTGTGCTGGGAA	25460
Dp	2336	AGGAGAAAGGAGCATTAAGTTTACAGCCTTACAGAGCGTACACAATATCTGTGCTGGGAA	23959
Qy	2547	AACCAAGACATTTTATCTAATTTTTTATTTAATAGTTTGTGTCTTATCTTAAATAGA	26060
Dp	2396	AACCAAGACATTTTATCTAATTTTTTATTTAATAGTTTGTGTCTTATCTTAAATAGA	24555
Qy	2607	TTTAAATGTCAAACTGTAGACAAATATATATATTTATATATTTACAAATTTGACAAAAA	26666
Dp	2456	TTTAAATGTCAAACTGTAGACAAATATATATATTTATATATTTACAAATTTGACTAATAA	25155

RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AK001658	313 bp	linear	AK001658	AK001658.1	GI:7023050	oligo capping, fls (full insert sequence).	Homo sapiens (human)	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
									Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,

- 1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Makematsu, A., Haseguchi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohsayashi, M., Nishi, T., Shichihara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuna, M., Nishitani, A., Sudo, H., Hosoi, T., Kaku, Y., Kodaira, T., Kondo, H.

TITLE	JOURNAL PUBMED AUTHORS	REFERENCE
Swagwara,M., Takahashi,M.,Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihira,K., Katsuta,N., Saito,K., Tanikawa,M., Yamazaki,M., Nimomiya,K., Ieshibashi,T., Yanashita,H., MuraiKawa,K., Fujimori,K., Tanai,H., Kinata,M., Watanabe,M., Hiraoaka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Houta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Tateuchi,K., Arita,M., Imose,N., Mutsaers,K., Yuki,H., Oshima,A., Sasaki,N., Kotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hisiigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimoto,Y., Koniyama,M., Tashiro,H., Taniyama,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Osaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Ohtani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigetake,K., Senba,T., Matsumura,K., Nakaijima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togeishi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,H., Masuno,Y., Yamashita,R., Nagase,T., Nomura,Y., Kikuchi,H., Masuno,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004) 14702039	2	
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotoku,S., Yoshikawa,Y., Matsumura,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Sasaki,N. NEBO human cDNA sequencing project Unpublished 3 (bases 1 to 3138) Isogai,T. and Otsuki,T. Direct Submission Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail:genomic@helix.co.jp, Tel:81-338-52-3975, Fax:81-338-52-3986) NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	COMMENT	
FEATURES	SOURCE	
CDS		
Location/Qualifiers		
1..3138		
/organism="Homo sapiens"		
/mol_type="RNA"		
/db_xref="taxon:9606"		
/clone_id="NT2RP4000648"		
/cell_line="NT2"		
/cell_type="teratocarcinoma"		
/clone_lib="NT2RP4"		
/note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction." 65..1936 /note="unnamed protein product" /codon_start=1 /protein_id="BAA91817.1" /db_xref="GI:7023051" /translation="MERPEBKQSPPOPMGRLLRLGAERGPVHLIRKREWTIGRRR GCGLSPSNKLVSGDHRITLYDRKSQGYLTEDSTSGTVINKLVKKQRCPLOGV IIVYRNKEPNENVAIYYESLSQGQTQSFEMVPFCVAQAQIKLGSDDPTLASQ SIYTSKGGGGISPKVSGSEVASDEVSSFASALPDRTKYSFSLLEPDQDLEPERKK MRGEGIDLDNLGLVLVAQPRNAQTVHEVDVAAGKDPMETELTCIIICDLADCVSL OPGHNFACCACTGAMERSLTPCRCPVERICNNILINLVAYAILIOHPDKSRSED YROMADANKITQOMLOPKYVRPSDESGEDLISDVNSSESDISAPVYRCQCE TRQAOAQPHCPAPBGBBPAPQALGDAPPISVTSTTRAVDYVTEPLQSGNALCTCCPP MPRRAREDDPVPAQQCAVCCLQPFCHLTWLGCTRGTCCVLAIPCELINADCKCLDGV		







Db 1918 GACAAAGTTCAAAAACCTAAGCATCCAGAGCCCTGAGACGTTTCAGACCTGAGAGGTGAA 1977  
 QY 2127 GAGAGGCTGTTTAAATACAGACAGACAGTCAAGGTGTTTCAACGCCCCCTGAG 2186  
 Db 1978 GAGAGGCTGTTTAAATACAGACAGACAGTCAAGGTGTTTCAACGCCCCCTGAG 2037  
 QY 2187 GGAAGGACGAGAGGCTCTCCGACAGAGTCTCTGGGGTGACTCTTCTGTGAGACTTTTAC 2246  
 Db 2038 GGAAGGACGAGAGGCTCTCCGACAGAGTCTCTGGGGTGACTCTTCTGTGAGAGC--TTTAC 2095  
 QY 2247 CCTCTGAGTGAACCTCTCCGACAGGCCCCGGGGGCGCGAGCCCTCTGTGTGAGCGC 2306  
 Db 2096 CCTCTGAGTGAACCTCTCCGACAGGCCCCGGGGGCGCGAGCCCTCTGTGTGAGCGC 2155  
 QY 2307 TGGGACAGGCTCTGTGTGAGTGAACAGACAGACAGAGCTTCTGTGAACATGCGCGCG 2366  
 Db 2156 TGGGACAGGCTCTGTGTGAGTGAACAGACAGACAGAGCTTCTGTGAACATGCGCGCG 2215  
 QY 2367 TCCCGCGAGAGGAGGAGTTTGCTCTTTGTACATTTTCCGAACTACAGTTAAAGAG 2426  
 Db 2216 TCTGCGCAGAGGAGGAGTTTGCTCTTTGTACATTTTCCGAACTACAGTTAAAGAG 2275  
 QY 2427 AAGTCTGTTTTCAGAAAGTTTCAAGGAGAGGCAAGTTTATCAAAAACATTTGTTT 2486  
 Db 2276 AAGTCTGTTTTCAGAAAGTTTCAAGGAGAGGCAAGTTTATCAAAAACATTTGTTT 2335  
 QY 2487 AGGAGAGGAGAGTAAAGTTTACAGCTTACAGACGTAACAATATCTCTGCTGTGGAA 2546  
 Db 2336 AGGAGAGGAGAGTAAAGTTTACAGCTTACAGACGTAACAATATCTCTGCTGTGGAA 2395  
 QY 2547 AACCAAGAGTTTATCTATTTTATTTTATTTATAGGTTTGCTTATCTTATATAGA 2606  
 Db 2396 AACCAAGAGTTTATCTATTTTATTTTATTTATAGGTTTGCTTATCTTATATAGA 2455  
 QY 2607 TTTAAATGTCAAACTGTAGCAAAATATATATTTATTTATTTTCAAAATGACAAAA 2666  
 Db 2456 TTTAAATGTCAAACTGTAGCAAAATATATATTTATTTATTTTCAAAATGACAAAA 2515

RESULT 9  
 AX405642 3181 bp DNA linear PAT 14-JUN-2002  
 LOCUS AX405642  
 DEFINITION Sequence 57 from Patent WO0222660.  
 ACCESSION AX405642  
 VERSION AX405642.1 GI:21438723  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS 1  
 TITLE Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,  
 Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.  
 JOURNAL Novel nucleic acids and polypeptides  
 Patent: WO 0222660-A 57 21-MAR-2002;  
 HYSEQ, INC. (US)  
 FEATURES  
 source Location/Qualifiers  
 1..3181  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 96..1967  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD34749.1"  
 /db\_xref="GI:21438724"  
 /translation="MERPEKOSPPPOPMGRLLRLGAESEPHVLLKREWTIGRR  
 /CDLSPSKIVSGDHCRIVNDEKSGOVLDPSTSGYINTLKVXKCTPCLOGCV  
 IYLYVRKNEPBNVAYLYESLSKQKOTBSPBMVPCYAGKGLKLSGSDPTLASQ  
 SIVTSGGGGSLPKSGPSVASDEVSSFASALPDRKLTFSLEPQDDELDPEVKK  
 MGKGDGLDNGOLVAQPRRNQVTHEVDAAGKPKDEKLELTCTICIQDLDHCVSL  
 QPCMTFFCAACYSGMERSLSCTCRCPYERICNHLNNLVAYLIIQHPDKRSBED  
 VQSMDBRNKITQDMLQPKVRSFSPDESSDLBELSDVDSISSDISGYYVVCROCP

## ORIGIN

Query Match 83.8%; Score 2244.6; DB 6; Length 3181;  
 Best Local Similarity 92.3%; Pred. No. 0;  
 Matches 2447; Conservative 0; Mismatches 79; Indels 125; Gaps 3;

YRROAOPPHCPAPEBPEAPALDADAPTSVLTAAVODYCPHQSHALCTCCFOP  
 MPDRABEDPRVAPPOCAVLOPCHLYMCSTRGCGCLAPFCMLGDKLDGV  
 LNNNSYEDILKNVYATRGTLTKMLTSVLAORGVELLSYRYTGDPVLCYCCGLR  
 SFRELTYQIQNTIPASELPVAVTSRDCWGNCRTOYKAHHAMKFNHICEOTFRN"  
 16 GGCCTCAATGTCTTTGACAGGCGCGCGGCGAGCCGGTTCGGGCGGGGCG 75.  
 Db 21 GGCCTCAATGTCTTTGACAGGCGCGCGGCGAGCCGGTTCGGGCGGGGCG 80  
 QY 76 GGAATGTGAATCCCATGAGAGGCGCCGAGAAAGCAACATCCCGCGCGAGGCC 135  
 Db 81 GGAATGTGAATCCCATGAGAGGCGCCGAGAAAGCAACATCCCGCGCGAGGCC 140  
 QY 136 TGGGACGAGCTCTGTGTGAGTGAACAGACAGAGGAGGAGCGGACGTCCTCTGAGAG 195  
 Db 141 TGGGACGAGCTCTGTGTGAGTGAACAGACAGAGGAGGAGCGGACGTCCTCTGAGAG 200  
 QY 196 CGGAGGTGACCATCGGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 255  
 Db 201 CGGAGGTGACCATCGGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 260  
 QY 256 GTCTGTGAGATCATCTGTGAATTTGTAGTGAATGAATAATCAGTGCAGGTGACATGGA 315  
 Db 261 GTCTGTGAGATCATCTGTGAATTTGTAGTGAATGAATAATCAGTGCAGGTGACATGGA 320  
 QY 316 GATACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 375  
 Db 321 GATACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 380  
 QY 376 CCTTTACAGACTGGGAGTGCATCTTGTGTGAACAGAGATGAACCGGAAACAAC 435  
 Db 381 CCTTTACAGACTGGGAGTGCATCTTGTGTGAACAGAGATGAACCGGAAACAAC 440  
 QY 436 GTGGCTACCTCTTATGAATCTTTAAGTGAAGCAAGGATGACACAGAAATCTTTGA 495  
 Db 441 GTGGCTACCTCTTATGAATCTTTAAGTGAAGCAAGGATGACACAGAAATCTTTGA 500  
 QY 496 GCTAACAGAGAAATGTGTTCCATGAGGACCAAGAATACCTCAGAGTGCAGAGGCGA 555  
 Db 501 ATGTGCTCTGT 551  
 QY 556 GGGGCGGATCCCGGGTCCCTCGTGTGCGCCGCGCATCAGGTGTCTTTGAGAACCA 615  
 Db 552 -----GATCTTCCCATCTTGCGCTCCCAAGTATGTGATTACA----- 590  
 QY 616 CAGCATCAACATCGACGTCAAGCTCTTCCCAAGCTCGGCTCTTCCACGAGGCT 675  
 Db 591 ----- 590  
 QY 676 TCTCTGACAGGCGAGAGGCTTCTTCAGATTGNGGCTGGGGGTGTGAGCATTCCTCT 735  
 Db 591 -----GGGTCTGGGGGTGTGAGCATTCCTCTCT 617  
 QY 736 AAGAGAGTGTCTCTCTGTGCAAGTGAATGATCTCCAGCTTTCCTCAGCTCTCCA 795  
 Db 618 AAGAGAGTGTCTCTCTGTGCAAGTGAATGATCTCCAGCTTTCCTCAGCTCTCCA 677  
 QY 796 GACAGAAAGCTGCGCTCTTTTGTGTGGAACCCGAGATCAGAGGATTTGAGGCC 855  
 Db 678 GACAGAAAGCTGCGCTCTTTTGTGTGGAACCCGAGATCAGAGGATTTGAGGCC 737  
 QY 856 GTGAAGAAAGTGAAGAGAGATGAGGAGCTTTGACCTGGAAGGAGGAGTGTGTGAGCA 915  
 Db 738 GTGAAGAAAGTGAAGAGAGATGAGGAGCTTTGACCTGGAAGGAGGAGTGTGTGAGCA 797  
 QY 916 CAACCGCTGAGAAATGCCCAACCGTCCACGAGACGTCAGAGCAGCGGCTGGAGAGCA 975  
 Db 798 CAACCGCTGAGAAATGCCCAACCGTCCACGAGACGTCAGAGCAGCGGCTGGAGAGCA 857

QY 976 GACAAATGAGAGAGAGCTGACATCATCTGCGAGAGCTGTCAGACGATCTGCG 1035  
 DB 858 GACAAATGAGAGAGAGCTGACATCATCTGCGAGAGCTGTCAGACGATCTGCG 917  
 QY 1036 AGTTTGAAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 1095  
 DB 918 AGTTTGAAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 977  
 QY 1096 TCGTCCCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1155  
 DB 978 TCGTCCCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037  
 QY 1156 AACAACTGT 1215  
 DB 1038 AACAACTGT 1097  
 QY 1216 GTGCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275  
 DB 1098 GTGCAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1157  
 QY 1276 CGGTCTTTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1335  
 DB 1158 CGGTCTTTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1217  
 QY 1336 AGTGAATCTCTGAGACATTTAGCCAGCATACGTGTGTGTGTGTGTGTGTGTGT 1395  
 DB 1218 AGTGAATCTCTGAGACATTTAGCCAGCATACGTGTGTGTGTGTGTGTGTGTGT 1277  
 QY 1396 AGGCAAGGCGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1455  
 DB 1278 AGGCAAGGCGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1337  
 QY 1456 GCGCTGAG 1515  
 DB 1338 GCGCTGAG 1397  
 QY 1516 TGGCCCTGT 1575  
 DB 1398 TGGCCCTGT 1457  
 QY 1576 AGAGCGAG 1635  
 DB 1458 AGAGCGAG 1517  
 QY 1636 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1695  
 DB 1518 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1577  
 QY 1696 TTTTGTGAGCTCAACCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755  
 DB 1578 TTTTGTGAGCTCAACCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1637  
 QY 1756 GAGTCAGACATCTGAGAGATTAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815  
 DB 1638 GAGTCAGACATCTGAGAGATTAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697  
 QY 1816 ACCGAG 1875  
 DB 1698 ACCGAG 1757  
 QY 1876 GAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1935  
 DB 1758 GAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1817  
 QY 1936 TATGCGAGAGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1995  
 DB 1818 TATGCGAGAGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1877  
 QY 1996 TATGCGAGAGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2055  
 DB 1878 TATGCGAGAGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1937

QY 2056 ATCTGTGAACAGACAG 2115  
 DB 1938 ATCTGTGAACAGACAG 1997  
 QY 2116 CTGAG 2175  
 DB 1998 CTGAG 2057  
 QY 2176 AGCCCTGTGAG 2235  
 DB 2058 AGCCCTGTGAG 2117  
 QY 2236 GAGCTTTTACCTCTGAGTGAACCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2295  
 DB 2118 GAGC--TTTACCTCTGAGTGAACCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2175  
 QY 2236 CTGAG 2355  
 DB 2176 CTGAG 2235  
 QY 2356 ACATGCGAG 2415  
 DB 2236 ACATGCGAG 2295  
 QY 2416 AGTTAAAGCAG 2475  
 DB 2296 AGTTAAAGCAG 2355  
 QY 2476 AACATGTTTACAG 2535  
 DB 2356 AACATGTTTACAG 2415  
 QY 2536 GCTGCTGAG 2595  
 DB 2416 GCTGCTGAG 2475  
 QY 2596 TTCTTAATTAAGTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2655  
 DB 2476 TTCTTAATTAAGTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2535  
 QY 2656 ATTGACAAAA 2666  
 DB 2536 ATTGACAAAA 2546

RESULT 10  
 AX834949 2448 bp DNA 1linear PAT 15-DEC-2003  
 LOCUS  
 DEFINITION Sequence 2073 from Patent EP1347046.  
 ACCESSION AX834949  
 VERSION AX834949.1 GI:39921084  
 KEYWORDS  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 Isegai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,  
 Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,  
 Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and  
 Masuko,Y.  
 TITLE Full-length cDNA sequences  
 JOURNAL Patent: EP 1347046-A 2073 24-SEP-2003;  
 FEATURES location/Qualifiers  
 source 1..2448  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 71.6%; Score 1919.4; DB 6; Length 2448;  
 Best Local Similarity 88.5%; Pred. No. 0;

Matches 2227; Conservative 0; Mismatches 6; Indels 284; Gaps 2;			
QY	218	GACGAGTTGCCGACCTTTCTCTTCCCGACAAATACTGCTCTTGGAGATCACTGTAGAA	277
DB	83	GACCTGTTGCCGACCTTTCTCTTCCCGACAAATACTGCTCTTGGAGATCACTGTAGAA	142
QY	278	TTGTAGTGAATAAATCAGGTCAAGTGAACACTGGAAATACAGACAAGTGAACAG	337
DB	143	TTGTAGTGAATAAATCAGGTCAAGTGAACACTGGAAATACAGACAAGTGAACAG	202
QY	338	TGATTAACAAGCTGAAGTTGTTAAGAAAGACAATGCTTTTACAGCTGGGATGTCA	397
DB	203	TGATTAACAAGCTGAAGTTGTTAAGAAAGACAATGCTTTTACAGCTGGGATGTCA	262
QY	398	TCTACTTGGTGTACAGAAAGATGAACCGGAAACAACAGTGGCATACCTCTATGAATCTT	457
DB	263	TCTACTTGGTGTACAGAAAGATGAACCGGAAACAACAGTGGCATACCTCTATGAATCTT	322
QY	458	TAAAGTAAAGCAAGGCGATGACACAGAAATCTTTAGACTAACAGAAATGTGTCC	517
DB	323	TAAAGTAAAGCAAGGCGATGACACAGAAATCTTTAGACTAACAGAAATGTGTCC	356
QY	518	ATGGGACAAAGATACCTCAGGTGACAGTGCAGGCGAGGCGCATCCCGGCTCCCTC	577
DB	357	-----	356
QY	578	CGTCTGCGCCCGCACTCAGGTGTGCTTTGAGGAACCAAGCCATCAATGCAAGCTCAG	637
DB	357	-----	356
QY	638	ACCTTTCGCCACAGCGCTCGGCTCTTCCAAGAGCTTCTCTGCAAGGCGAGAGCGTT	697
DB	357	-----	356
QY	698	CCTCCAGTTGTGGTGTGGGGGTGTGTGGAATCTCCCTAAAGAGATGTCCCTCTGTG	757
DB	357	-----TGGGTCTGGGGGTGTGTGGAATCTCCCTAAAGAGATGTCCCTCTGTG	406
QY	758	CAAGTATGAAGTCTCAGCTTTGCTCAGCTTCTCCCAACAGAAAGACTGCGTCTTTT	817
DB	407	CAAGTATGAAGTCTCAGCTTTGCTCAGCTTCTCCCAACAGAAAGACTGCGTCTTTT	466
QY	818	CGTGTGGAAACCCGAGATCAGAGAGATTTGGAGCCGCTGAAGAAATGAGAGAG	877
DB	467	CGTGTGGAAACCCGAGATCAGAGAGATTTGGAGCCGCTGAAGAAATGAGAGAG	526
QY	878	ATGGGACCTTGAACCTGAACGCGCAGTTGTGTGCAACAACCGCTAGAAATGCCCAA	937
DB	527	ATGGGACCTTGAACCTGAACGCGCAGTTGTGTGCAACAACCGCTAGAAATGCCCAA	586
QY	938	CGTCCACAGAGAGCTCAAGAGAGCGGCTGGGAAGCCAGACAAAGATGAGAGAGCGCTGA	997
DB	587	CGTCCACAGAGAGCTCAAGAGAGCGGCTGGGAAGCCAGACAAAGATGAGAGAGCGCTGA	646
QY	998	CATGCAATCTGCGCAGAGACTGCTGCAAGACTGCGTGAAGTTGAGCCCTGATGACACA	1057
DB	647	CATGCAATCTGCGCAGAGACTGCTGCAAGACTGCGTGAAGTTGAGCCCTGATGACACA	706
QY	1058	CGTTCGCGCGCTTGTCTACTCGGCGTGAATGAGGCGTCTCTCTGTGCTTCACTGTGC	1117
DB	707	CGTTCGCGCGCTTGTCTACTCGGCGTGAATGAGGCGTCTCTCTGTGCTTCACTGTGC	766
QY	1118	GCTGTCCCGTGAAGCGGATCTGTAAACAATCTCTCAACAACCTGTGGAAGCAATACC	1177
DB	767	GCTGTCCCGTGAAGCGGATCTGTAAACAATCTCTCAACAACCTGTGGAAGCAATACC	826
QY	1178	TCATCAGAGATCCAGCAAGAGTGCAGTGAAGAAGATGCAAAATATGAGATGCCAGA	1237
DB	827	TCATCAGAGATCCAGCAAGAGTGCAGTGAAGAAGATGCAAAATATGAGATGCCAGA	886
QY	1238	ATTAATACTCACTCAAGACATGCTGCAAGCCAAAGTCAAGCGGTCTTTTCTGTAGAGAAG	1297
DB	887	ATTAATACTCACTCAAGACATGCTGCAAGCCAAAGTCAAGCGGTCTTTTCTGTAGAGAAG	946

QY	1298	GGAGTTTCAAGAGACCTGCTGAGCTGTCAAGAGTTGAACAGTAGTCTCAGACATTAAGCC	1357
DB	947	GGAGTTTCAAGAGACCTGCTGAGCTGTCAAGAGTTGAACAGTAGTCTCAGACATTAAGCC	1006
QY	1358	AAGCAATACGTCTGTGTGCGCGCAGATGCTCTGATTAAGAAAGAGCGCGCAGCTCCCC	1417
DB	1007	AAGCAATACGTCTGTGTGCGCGCAGATGCTCTGATTAAGAAAGAGCGCGCAGCTCCCC	1066
QY	1418	ACTGCCAGACACCCAGAGGCGAGCCAGAGCCCAAGGCCCTTGAGGATGCACTCTCA	1477
DB	1067	ACTGCCAGACACCCAGAGGCGAGCCAGAGCCCAAGGCCCTTGAGGATGCACTCTCA	1126
QY	1478	CGTCCGTGACCTTGAAGACAGAGATGCAAGATTAAGTGTGCCCTCTGCAAGGAAGCAGC	1537
DB	1127	CGTCCGTGACCTTGAAGACAGAGATGCAAGATTAAGTGTGCCCTCTGCAAGGAAGCAGC	1186
QY	1538	CCCTGTGACCTGTGCTTCCAGCCCATGCGGACCGGAGAGCGGACCGAGAGAGACC	1597
DB	1187	CCCTGTGACCTGTGCTTCCAGCCCATGCGGACCGGAGAGCGGACCGAGAGAGACC	1246
QY	1598	CGCGTGTGCGCCCTTCAAGAGTGTGCGGTGTGCTGCAAGCTTTTGTGCACTGTGAG	1657
DB	1247	CGCGTGTGCGCCCTTCAAGAGTGTGCGGTGTGCTGCAAGCTTTTGTGCACTGTGAG	1306
QY	1658	GCTGCAACCCGAGACCGGCTGCTACGGCTGCTGCGGCGCTTTTGT-----	1701
DB	1307	GCTGCAACCCGAGACCGGCTGCTACGGCTGCTGCGGCGCTTTTGT-----	1366
QY	1702	-----GAGCTCAA	1709
DB	1367	GCCTTCGTGCGACATCATTTAATAAGTAAAGTAACTTGCCTTCACTCGGACAGAGCTCAA	1426
QY	1710	CCTGGGTGAACAAGTGTCTGGAACGCGGTGTGTAACAACAACAGCTACAGATCAATCTT	1769
DB	1427	CCTGGGTGAACAAGTGTCTGGAACGCGGTGTGTAACAACAACAGCTACAGATCAATCTT	1486
QY	1770	GAAAGATTAACCTGGGAAACAGAGATTTGAACAGAAACAATGTAACCGAAGAGCTGCT	1829
DB	1487	GAAAGATTAACCTGGGAAACAGAGATTTGAACAGAAACAATGTAACCGAAGAGCTGCT	1546
QY	1830	GGCTCTCAGCGGGAGTGTCTTCTGCTGTCTGATTAACAGATCAAGGAGACACGTTCT	1889
DB	1547	GGCTCTCAGCGGGAGTGTCTTCTGCTGTCTGATTAACAGATCAAGGAGACACGTTCT	1606
QY	1890	GTGTTAAGTGTGTGCTGTGCGGAGCTTCCGTAGCTGACCTATCAGTATCGGCAAGACAT	1949
DB	1607	GTGTTAAGTGTGTGCTGTGCGGAGCTTCCGTAGCTGACCTATCAGTATCGGCAAGACAT	1666
QY	1950	TCCTGTCCGAGTTGCGCAGTGTGCGGTAAACATCCCGTCTGTACTGTAGCGGCGGTAA	2009
DB	1667	TCCTGTCCGAGTTGCGCAGTGTGCGGTAAACATCCCGTCTGTACTGTAGCGGCGGTAA	1726
QY	2010	CTGCCGACTCAGGTGAAGACTCAACAAGCACTGAATTAATCAATCATCTGTGAACAGAC	2069
DB	1727	CTGCCGACTCAGGTGAAGACTCAACAAGCACTGAATTAATCAATCATCTGTGAACAGAC	1786
QY	2070	AAAGTTCAAAAATTAAGCAATCCAGAGCCCTGAGAGCTTTTCAAGACTGAGGTGAAGAG	2129
DB	1787	AAAGTTCAAAAATTAAGCAATCCAGAGCCCTGAGAGCTTTTCAAGACTGAGGTGAAGAG	1846
QY	2130	AAGGTGTTTTTAATAATCAAGAGACAGTCAAGGTGTTTTTCAAGCGCCCTGAGAGGA	2189
DB	1847	AAGGTGTTTTTAATAATCAAGAGACAGTCAAGGTGTTTTTCAAGCGCCCTGAGAGGA	1906
QY	2190	AAGGACGAGGCTCTCCGACAGGTCTCTGAGGTGACTTTCTGTGAGACTTTTAACCT	2249
DB	1907	AAGGACGAGGCTCTCCGACAGGTCTCTGAGGTGACTTTCTGTGAGACTTTTAACCT	1966
QY	2250	CTGAGTGAAGACCTTCCCAAGACCCCGGAGGCGGAGCCCGCTCTGTGAGGCGCTGG	2309
DB	1967	CTGAGTGAAGACCTTCCCAAGACCCCGGAGGCGGAGCCCGCTCTGTGAGGCGCTGG	2026



QY 758 CAAGTATGAAAGTCTCAGCTTGGCTCAGCTCTCCAGACAGAAAAGACTGCGTCTTTT 817  
 DB 407 CAAGTATGAAAGTCTCAGCTTGGCTCAGCTCTCCAGACAGAAAAGACTGCGTCTTTT 466  
 QY 818 CGTCTTGGAAACCCAGAGATCAGAGAGATTGGAGCCCGTGAAGAAATAGAGAGAG 877  
 DB 467 CGTCTTGGAAACCCAGAGATCAGAGAGATTGGAGCCCGTGAAGAAATAGAGAGAG 526  
 QY 878 ATGGGGAACCTTGAACCTTGAAGAGGCAATTGGTCCGCAAAACCGCTGAAGAAATGCCCAA 937  
 DB 527 ATGGGGAACCTTGAACCTTGAAGAGGCAATTGGTCCGCAAAACCGCTGAAGAAATGCCCAA 586  
 QY 938 CCGTCCAGAGAGAGCTCAGAGAGCGGCTGGAGAGCAGACAGATGAGAGAGAGCTGA 997  
 DB 587 CCGTCCAGAGAGAGCTCAGAGAGCGGCTGGAGAGCAGACAGATGAGAGAGAGCTGA 646  
 QY 998 CATGCATCATCTTGCAGAGACCTTGCCTGCACTGCGTGAATTGGACCCCTGCATGACAA 1057  
 DB 647 CATGCATCATCTTGCAGAGACCTTGCCTGCACTGCGTGAATTGGACCCCTGCATGACAA 706  
 QY 1058 CGTCTTGGCGGCTTGTCTACTCGGGCTGCAATGAGAGGCTGTCTCTGTCTTACTGCTGCC 1117  
 DB 707 CGTCTTGGCGGCTTGTCTACTCGGGCTGCAATGAGAGGCTGTCTCTGTCTTACTGCTGCC 766  
 QY 1118 GCTGTCCCGTGGAGCGGATCTGTAAACCAATCCCTCAACAACTCGTGGAGAGCAATACC 1177  
 DB 767 GCTGTCCCGTGGAGCGGATCTGTAAACCAATCCCTCAACAACTCGTGGAGAGCAATACC 826  
 QY 1178 TCATCCAGATCCAGACAAAGAGTCGAGTGAAGAGATGTGCAAAATGATGATGCAAGAA 1237  
 DB 827 TCATCCAGATCCAGACAAAGAGTCGAGTGAAGAGATGTGCAAAATGATGATGCAAGAA 886  
 QY 1238 ATTAATCACTCAAGACATGCTGCAAGCCAAAGTCAGGCGGCTTTTCTGATGAAGAG 1297  
 DB 887 ATTAATCACTCAAGACATGCTGCAAGCCAAAGTCAGGCGGCTTTTCTGATGAAGAG 946  
 QY 1298 GGAGTTCAGAGAGAGCTGTGAGAGCTGTGACAGCTTGAACAGTGAAGCTCTCAGACATTAAGC 1357  
 DB 947 GGAGTTCAGAGAGAGCTGTGAGAGCTGTGACAGCTTGAACAGTGAAGCTCTCAGACATTAAGC 1006  
 QY 1358 AGCCATACATCTGTGTGCTGCTGAGTCTGATGACAGAAAGAGAGGCGGCGCAGCTCTCCC 1417  
 DB 1007 AGCCATACATCTGTGTGCTGCTGAGTCTGATGACAGAAAGAGAGGCGGCGCAGCTCTCCC 1066  
 QY 1418 ACTGCCAGAGACCCAGAGGAGAGCCAGAGAGCCCAAGGCTCTGGAGAGATGACCTTCCA 1477  
 DB 1067 ACTGCCAGAGACCCAGAGGAGAGCCAGAGAGCCCAAGGCTCTGGAGAGATGACCTTCCA 1126  
 QY 1478 CGTCCGTCAAGCTGAGAGAGAGTCAAGATTAAGTGTGCTCTGCAAGAGAGAGCAGAG 1537  
 DB 1127 CGTCCGTCAAGCTGAGAGAGAGTCAAGATTAAGTGTGCTCTGCAAGAGAGAGCAGAG 1186  
 QY 1538 CCGTGTGACATCTGCTCTTCAAGCCATGCGCCAGAGCGAGAGCGCAGAGCAGAGAGC 1597  
 DB 1187 CCGTGTGACATCTGCTCTTCAAGCCATGCGCCAGAGCGAGAGCGCAGAGCAGAGAGC 1246  
 QY 1598 CCGGTGTGCGCCCTCAAGAGTGTGCGGTGCTGCAAGCTTCTGCAACTGTACTGAG 1657  
 DB 1247 CCGGTGTGCGCCCTCAAGAGTGTGCGGTGCTGCAAGCTTCTGCAACTGTACTGAG 1306  
 QY 1658 GGTGACCCCGGAGCCGCTGCAAGCTGAGCTGAGCCCGCTTTTGT 1701  
 DB 1307 GGTGACCCCGGAGCCGCTGCAAGCTGAGCCCGCTTTTGTGCTGAGAGAGAGT 1366  
 QY 1702 -----GAGCTCAA 1709  
 DB 1367 GCTTCTGAGACATGATTAATAAGATTAACATTTGCCCTTCACTCGGAGAGCTCAA 1426  
 QY 1710 CCTGGGTGAACAGTGTCTGAGCGGCTGTGAACAACAACAGTCAAGTCAAGTCAATCCT 1769  
 DB 1427 CCTGGGTGAACAGTGTCTGAGCGGCTGTGAACAACAACAGTCAAGTCAAGTCAATCCT 1486

QY 1770 GAAGATTACTTGCAACACAGAGGTTTGCATGAGAAAAATGTTACGAGAGCTCGT 1829  
 DB 1487 GAAGATTACTTGCAACACAGAGGTTTGCATGAGAAAAATGTTACGAGAGCTCGT 1546  
 QY 1830 GGCTTCCAGCGGAGAGTGTCTGTCTGATTAACAAGTCAAGAGAGACCGTCTCT 1889  
 DB 1547 GGCTTCCAGCGGAGAGTGTCTGTCTGATTAACAAGTCAAGAGAGACCGTCTCT 1606  
 QY 1890 GTGTATCTGTGTGCTGCGAGAGCTTCCGAGAGCTGACCTATTCGAGAGAACT 1949  
 DB 1607 GTGTATCTGTGTGCTGCGAGAGCTTCCGAGAGCTGACCTATTCGAGAGAACT 1666  
 QY 1950 TCTGTCTTCCAGATTGCGAGTGCCTGAACATCCCGTCTGACTGAGGCGCTGA 2009  
 DB 1667 TCTGTCTTCCAGATTGCGAGTGCCTGAACATCCCGTCTGACTGAGGCGCTGA 1726  
 QY 2010 CTGCGCACTCAGGTGAAGGCTCACAGCGCATGAATTAATCATATCTGTGAACAGAC 2069  
 DB 1727 CTGCGCACTCAGGTGAAGGCTCACAGCGCATGAATTAATCATATCTGTGAACAGAC 1786  
 QY 2070 AAGGTTCAAAAATTAAGCATCCAGAGAGCCCTGAGAGCTTTGACAGCTGAGAGTGAAG 2129  
 DB 1787 AAGGTTCAAAAATTAAGCATCCAGAGAGCCCTGAGAGCTTTGACAGCTGAGAGTGAAG 1846  
 QY 2130 AAGGTTTAAAAATTAAGAGACAGACGTCAGAGTGTGTTCAAGGCCCTGAGAG 2189  
 DB 1847 AAGGTTTAAAAATTAAGAGACAGACGTCAGAGTGTGTTCAAGGCCCTGAGAG 1906  
 QY 2190 AAGGAGCAGAGGCTCTCCAGACAGTGTCTGAGGAGTCTTCTGTGAGACTTTTACCT 2249  
 DB 1907 AAGGAGCAGAGGCTCTCCAGACAGTGTCTGAGGAGTCTTCTGTGAGACTTTTACCT 1966  
 QY 2250 CTGAGTGAACCTCTCCAGAGAGCCCGGAGGCGGAGCCCGGCTCTGTGAGAGCTG 2309  
 DB 1967 CTGAGTGAACCTCTCCAGAGAGCCCGGAGGCGGAGCCCGGCTCTGTGAGAGCTG 2026  
 QY 2310 GCAGGAGCTGTGTGATCAGACAGAGAGAGCAAGCTTCTGTGAACAGGCGGCTG 2369  
 DB 2027 GCAGGAGCTGTGTGATCAGACAGAGAGAGCAAGCTTCTGTGAACAGGCGGCTG 2086  
 QY 2370 CGCCGAGAGGAGAGTGTGCTCTTGTGACATTTTCGAAACTACAGTTAAAGAGAG 2429  
 DB 2087 CGCCGAGAGGAGAGTGTGCTCTTGTGACATTTTCGAAACTACAGTTAAAGAGAG 2146  
 QY 2430 TCTGTTTTCAGAGAAAGTTTCAAGAGAGAGGCAAGTTTATCAAAAATGTTTCAAG 2489  
 DB 2147 TCTGTTTTCAGAGAAAGTTTCAAGAGAGAGGCAAGTTTATCAAAAATGTTTCAAG 2206  
 QY 2490 AGAAGGAGCATTAAGTTTACAGCTTACAGAGAGTACAAATATCTGTGAGAGAG 2549  
 DB 2207 AGAAGGAGCATTAAGTTTACAGCTTACAGAGAGTACAAATATCTGTGAGAGAG 2266  
 QY 2550 CACAGCATTTTATTAATTTTATTAATTAAGTGTGCTTATCTTAATTAAGATT 2609  
 DB 2267 CACAGCATTTTATTAATTTTATTAATTAAGTGTGCTTATCTTAATTAAGATT 2326  
 QY 2610 AATGTCAAAACTGTAGACAAATATTAATTTATTAATTTAACAATTTGACAAA 2666  
 DB 2327 AATGTCAAAACTGTAGACAAATATTAATTTATTAATTTAACAATTTGACAAA 2383

RESULT 12  
 BC049792 3194 bp mRNA linear ROD 30-JUN-2004  
 LOCUS BC049792  
 DEFINITION Mus musculus checkpoint with forkhead and ring finger domains, mRNA  
 (cDNA clone MGC:59419 IMAGE:6510480), complete cds.  
 ACCESSION BC049792  
 VERSION BC049792.1 GI:29437110  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.







```

Db      730  ACCTGTGCAATGGGGACTGCTAGCCCTTCTCCAGTTTCCAGACAAAGA---ACG 786
Qy      810  GTTCCTTTTGTGCTGGAAACCCAGGATCAAGAGATTGGAGCCCGTGAAGAAAT 869
Db      787  ATCTCTTTTGTGCTGGAAAGTAAACCATGAGGAAATGGAGCCCTGCCAAAAAAGAT 846
Qy      870  GAGAGAGATGGGGAACCTTGACCTGAACGGGGAGTTGGTGTGCGACAAACGGGTAGAA 929
Db      847  GAAAGAGATGGGGAACCTTGACAGAACCTCCAGTTATTAATTTCAGGCCAGGCTGGAAA 906
Qy      930  TGCCCAAAACCGTCCACAGAGAGCTCAAGACAGCGGCTGGAAACCCAGACAAAGTGAAGA 989
Db      907  TGCCCAAAACCTCAAGAGAGATGTCAAAAGATGCTGTGTGAGGCCAGACAAAGTGAAGA 966
Qy      990  GACGCTGACATGATATATCTGCCAGAACTGTGCGACGACTGCGTAGTTTGACGCTTG 1049
Db      967  GACACTAACCTGATATCTGCGACAGAACCTTCTGCAAGATGTGTGAGTTTGCAGCCTTG 1026
Qy      1050  CATGCAACAGTTCTGCGGGGCTTGTACTCGGGCTGAGTGAAGCGCTGCTGCTGTGTC 1109
Db      1027  TATGCAACATTTTGTGCGGCTTGTACTCTGTGTGATGAGCGTTCAATCTGTGCCC 1086
Qy      1110  TACCTGCGCTGTCCCGTGAAGCGGATCTGTAAAAACCAATCTCAACACCTGCTGA 1169
Db      1087  TACCTGCGGATGTCAGTGAAGGAGATTGCAAAAACCAATCTGAAACAACCTAGTGA 1146
Qy      1170  AGCATACCTCATTCAGACATCCAGACAAAGTGGCGAGTGAAGAAAGATGTCAAGTAGA 1229
Db      1147  AGCATACCTTATTCAGACACCCAGATMAAGTCGAGTGAAGAAAGATGTCAAGTAGA 1206
Qy      1230  TGCCAGAAATAAATCACTCAAGACATGCTGACAGCCCAAGTCAGGCGCTTTTCTGA 1289
Db      1207  TGCAAGAAATAAATCACTCAAGATATGCTGCAACCAAGTCAAGAGGCTTTTCTGA 1266
Qy      1290  TGAAGAGGAGATTCAAGAGACCTGTGAGCTGTCAAGACTGTGAACATGATCTCTGA 1349
Db      1267  TGAAGAGGAGATTCAAGAGACCTGTGAGCTGTGAATGTGCAACATGATCTCTGA 1326
Qy      1350  CATTAAGCCAGCATATGCTGTGTGCGGAGTGTCTGTAGTACAGAAAGCAGGCGCGCA 1409
Db      1327  TATCATGTCAGCATATGCTGTGCAAGATGTCTGTAAATACAGAAAGCAGGCGTGA 1386
Qy      1410  GCTCCCAACCTGCCAGACCCGAGGGGAGCCAGGAGCCCAAGGCGCT---GGGGGA 1466
Db      1387  GTCTCTTCTTCCAGCTCCAGAGAGTAGCTGGAAGCTACATGCGCCCTTGTGAGGGA 1446
Qy      1467  TGCAACCTTCAGCTGCGTCAAGCTGACAGACATGTCAGAGTTAGTGTGCTCTGA 1526
Db      1447  GGCACCTTCAATCTGCGACGCTTGCAACAGACGCCAAGATTAATGTGCTCTCTGA 1506
Qy      1527  AGGAAGCCAGCGCTGTGTGCACTGTGCTTCAAGCCCATGCGCAACCGAGAGCGAGAG 1586
Db      1507  AGGAAGCCATGCAATATGCACTGTGCTTCAAGCCTATGCTGACCGGAGAGCTGAAG 1566
Qy      1587  CGAAGAGAACCCGCGTGTGCGCCCTCAAGAGATGTGTGCTGCTGCGACCTTTCTGCA 1646
Db      1567  GGAGAGAGATCCCGCGTGTGCGCCCTCAAGAGATGTGTGCTGCGACCTTTCTGCA 1626
Qy      1647  CCTGTACTGAGGCTGCAACCGGACCGGCTGCTACGGCTGCTGCGCCCGCTTTGTAGCT 1706
Db      1627  CCTGTACTGAGGCTGCAACGAGATCTGTGCTGCTTTGTGCTGCTGCGCCCATTTCTGAGCT 1686
Qy      1707  CAACCTGGGTGACAAAGTGTCTGACGCGGTGTGAACAACAACAGTACAGTACAGACAT 1766
Db      1687  CAACCTGGGTGACAAAGTGTCTGAGATGAGATGCTGAACAATTAATTAATTAATTAAT 1746
Qy      1767  CCGTAAGAAATTAATCTGAGCAACAGAGGTTTGAATGAGAAACAAATTTGACCGAAGCT 1826
Db      1747  CCGTAAGAAATTAATCTGAGCAACAGAGGCTTGAATGAGAAATGTGTGACAGAGAGCT 1806
Qy      1827  CGTGTCTCTCAAGCGGGAGTGTCTGTGCTGTGATTAACAGATGACGGAGACACCGT 1886

```

```

Db      1807  CCTGTCTGACGAGGATGATTAATTAATGCTGTGATTAACAGAAATCACTGAATAACTGT 1866
Qy      1887  TCTGTACTACTGCTGTGCTGCGGACGCTTCCGTGAGCTGACTATCACTATCCGACAA 1946
Db      1867  GCTGTGTACTGCTGTGCTGCTGAGTCTTCCGAAGCTGACCTTCAAGTATCTGACAA 1926
Qy      1947  CATTCCTGCTCCAGATTGCGAGTGCAGTACATCCGCTCTGACTGCTGAGGAGCG 2006
Db      1927  CATTCCTGCTTCTGAGTTGCGAGTACATTAATATCCGCTCTGACTGCTACTGAGGCG 1986
Qy      2007  TAACTGCGCACTAGTGAAGAGCTCACCAAGCCATGAATTAATCAATATCTGTGAACA 2066
Db      1987  CAATGTGCACTAGTGAAGAGCTCACCAATGAATTAATTAATCAATATCTGTGAACA 2046
Qy      2067  GACAGGTTCAAAATTAAGCATCCAGAGGCGCTGAGCAGCTTTCAAGCATGAGAGTGA 2126
Db      2047  AACAGGTTCAAGAACTGAATTTGAGAGAGACAGTGAACAAGTTACAGACACAGAG 2106
Qy      2127  GAGAGCGTGTTTTAAATTAACAGAGACAGCATCAAGGTGTTTCAAGCCCTGAG 2186
Db      2107  GCAT- GTTTTAAAGGCCCATTAACAACCTTGAAGGCAATTTTACATTTCCCT 2165
Qy      2187  GGAAGGACGCAAGGCTTCCCAACAGTGTCTGAGGAGTACTTCTGTGAGCTTTTAC 2246
Db      2166  ACGGACGACTGGGGGTTCTGATCAGGCACTGGGAGTGTGAATTTCCCTTGTGAGCATC 2225
Qy      2247  CCTGTGATGAGACCTTCCCAAGACCCCGGGGGCGGAGCCGCTCTGTGTAGCGC 2306
Db      2226  ATGTCAGTGAACAACAACAACGTCGCTCAGTAAAC-CTGGAGGCGCTCTGCGCAACC 2284
Qy      2307  TGGGAGGCGCTGTGTGTGCAATCAGACAGAG-----ACGAAGCC 2347
Db      2285  AGCTTTTATGCTGTGTGAAGAGATCAATGATGATGATTTCTCCCTCCCACTGAGAC 2344
Qy      2348  TTTCTGTAAACATGCGGCGCTCCCGCGAGAGGGGAGTTTGTCTTTTGTATCATTTCC 2407
Db      2345  GCTTCCCTCTGTGTGTGAGGAGCATGACAGTGCAGGTTGTGATCTTTTGTATCTTTT-- 2402
Qy      2408  GAAACTTACGTTTAAAGCAGAACTCTGTTTTCAGGAAAAAGTTTCAAGGGAAGGCAAGT 2467
Db      2403  ----CTACCTGTATCTTGAATTTGTTTAAATGGAAGCTCATGGAAGATTTGTACGC 2458
Qy      2468  TTATCAAAAACATGTTTCAAGAG--AAGGAGCATTAAGTTTACAGCCATCAGAGCT-A 2524
Db      2459  TGTGTCAAGCGATTTCTCGGAAGAAAGAAAGCATAGCTTTACAGCTCCAGGGGTACA 2518
Qy      2525  CACAATATCTGCTGCTGCGGAAACACAGCATTTTATCTATTTTATTAATAGTT 2584
Db      2519  TAGATATGCTCTGTCTGGAAGAACACAGCTTTTGTATATTTTATTAATTAAGTT 2578
Qy      2585  TGTGCTTATCTTCTTAATTAAGTTTAAATGTCACAAATCTGTAGCAAAATTAATTAAT 2644
Db      2579  TGTGCTTATCTTCTTAATTAAGTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2638
Qy      2645  ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2675
Db      2639  ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2669

RESULT 13
BD183291      6235 bp   DNA      linear   PAT 17-JUN-2003
LOCUS        BD183291
DEFINITION   Novel genes and proteins encoded by the genes.
ACCESSION    BD183291
VERSION      BD183291.1 GI:31875491
KEYWORDS     JP 2002345492-A/4.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 6235)
AUTHORS      Ohara, O., Nagase, T., and Nakajima, D.
TITLE        Novel genes and proteins encoded by the genes

```

JOURNAL Patent: JP 2002345492-A 4 03-DEC-2002;  
 KAZUSA DNA RESEARCH INSTITUTE  
 COMMENT OS Homo sapiens (human)  
 PN JP 2002345492-A/4  
 PD 03-DEC-2002  
 PF 26-FEB-2002 JP 2002049009  
 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA  
 PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,  
 A61P25/14,  
 A61P25/18, A61P35/00, C12N15/00, A61K37/02  
 PC Novel genes and proteins encoded by the genes FH Key  
 CC Location/Qualifiers  
 FT CDS (50).. (1414).  
 Location/Qualifiers  
 source 1..6235  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 47.8%; Score 1279.8; DB 6; Length 6235;  
 Best Local Similarity 99.5%; Pred. No. 4,5e-285;  
 Matches 1306; Conservative 0; Mismatches 2; Indels 5; Gaps 2;  
 Oy 1354 AGCCAGCCATAGCTGTGTGTCGGGAGCTGCTCTGAGTACAGAAAGGACGGCGGAGGCT 1413  
 Db 4305 ACCCGACCATAGCTGTGTGTCGGGAGCTGCTCTGAGTACAGAAAGGACGGCGGAGGCT 4364  
 Oy 1414 CCCCATGCGCCAGCACCAGAGGGCGAGCAGAGAGCCCAAGGCCCTGGGGGATGACACC 1473  
 Db 4365 CCCCATGCGCCAGCACCAGAGGGCGAGCAGAGAGCCCAAGGCCCTGGGGGATGACACC 4424  
 Oy 1474 TCCAGTCCGTAGAGCTGACAGCAGAGCTCCAGATTACGTGTGCTCTGCAAGAAAGC 1533  
 Db 4425 TCCAGTCCGTAGAGCTGACAGCAGAGCTCCAGATTACGTGTGCTCTGCAAGAAAGC 4481  
 Oy 1534 CAGCGCTGTGACCGTGTGCTTCCAGCCATGCCCGAGCCGAGAGCGGAGCGGAGAG 1593  
 Db 4482 CAGCGCTGTGACCGTGTGCTTCCAGCCATGCCCGAGCCGAGAGCGGAGCGGAGAG 4541  
 Oy 1594 GACCGCGTGTGACCGTGTGCTTCCAGCGATGTGCGTCTGCTGACAGCCTTCTGCGACCTGTAC 1653  
 Db 4542 GACCGCGTGTGACCGTGTGCTTCCAGCGATGTGCGTCTGCTGACAGCCTTCTGCGACCTGTAC 4601  
 Oy 1654 TGGGCGTGTGACCGCGGACCGGCTGTACAGGCTGTGCGTCTGCTGAGCTCAACTG 1713  
 Db 4602 TGGGCGTGTGACCGCGGACCGGCTGTACAGGCTGTGCGTCTGCTGAGCTCAACTG 4661  
 Oy 1714 GGTGACAGATGTGTGACCGGCGTGTGACCAACAGCTACAGTACAGATCTCTGAAG 1773  
 Db 4662 GGTGACAGATGTGTGACCGGCGTGTGACCAACAGCTACAGTACAGATCTCTGAAG 4721  
 Oy 1774 AATTACCTGGCAACAGAGGTTTGAATGAAAAAATGTTGACCGAGAGCCTGTGCT 1833  
 Db 4722 AATTACCTGGCAACAGAGGTTTGAATGAAAAAATGTTGACCGAGAGCCTGTGCT 4781  
 Oy 1834 CTCGACGGGGAGGATTTTCTGTGTCTGATTAACAAGTCAAGGAGACACGCTTGTGT 1893  
 Db 4782 CTCGACGGGGAGGATTTTCTGTGTCTGATTAACAAGTCAAGGAGACACGCTTGTGT 4841  
 Oy 1894 TACTCTGTGGCTGTGACGAGCTTCCGTGAGCTGACCTATACAGTATCGGACAGCAATTCT 1953  
 Db 4842 TACTCTGTGGCTGTGACGAGCTTCCGTGAGCTGACCTATACAGTATCGGACAGCAATTCT 4901  
 Oy 1954 GCTTCCGAGTGTGACGAGGCGCTTAATCCGCTCTGACTGTGAGGCGCGTAATGTC 2013  
 Db 4902 GCTTCCGAGTGTGACGAGGCGCTTAATCCGCTCTGACTGTGAGGCGCGTAATGTC 4961  
 Oy 2014 CGCACTCGAGTGAAGCTCAACAGCCGATGAATTAATCAATCAATCTGTGAACAGCAAG 2073  
 Db 4962 CGCACTCGAGTGAAGCTCAACAGCCGATGAATTAATCAATCAATCTGTGAACAGCAAG 5021  
 Oy 2074 TTCAAAAATAAGCATCCAGAGGCGCTTGAAGAGCTTGAAGAGTGAAGAGAGCG 2133

Db 5022 TTCAAAAATAAGCATCCAGAGGCGCTTGAAGAGCTTGAAGAGTGAAGAGAGCG 5081  
 Oy 2134 TGTTTTAAATATACAGACAGACAGCTCAAGATGTTTTCACAGGCCCTTGAAGAGAGG 2193  
 Db 5082 TGTTTTAAATATACAGACAGACAGCTCAAGATGTTTTCACAGGCCCTTGAAGAGAGG 5141  
 Oy 2194 ACGAGGGTCTCCAGACAGTGTCTGTGGGTGATCTTCTGTGAGAGCTTTTACCTCTGA 2253  
 Db 5142 ACGAGGGTCTCCAGACAGTGTCTGTGGGTGATCTTCTGTGAGAGC--TTTACCTCTGA 5199  
 Oy 2254 GTGAGACCTTCCAGAGGCGCGGGGCGGAGAGCCGCTCTGTGTGAGAGCGCTGGAGAG 2313  
 Db 5200 GTGAGACCTTCCAGAGGCGCGGGGCGGAGAGCCGCTCTGTGTGAGAGCGCTGGAGAG 5259  
 Oy 2314 GGTCTGTGTGATCAGACAGACAGAGAGAGCTTCTGTGAATGCGGCGCTCCGCC 2373  
 Db 5260 GGTCTGTGTGATCAGACAGACAGAGAGAGCTTCTGTGAATGCGGCGCTCCGCC 5319  
 Oy 2374 GAGAGGGGAGCTTGTCTCTTTGTACATTTTCCGAAATCTACAGTTAAAGAGAGTGTG 2433  
 Db 5320 GAGAGGGGAGCTTGTGTCTTTGTACATTTTCCGAAATCTACAGTTAAAGAGAGTGTG 5379  
 Oy 2434 TTTTCAAGGAAAGTTTCAAGGAGAGAGGCGAAGTTTTCAAAACATTTGTTTCAGAGAA 2493  
 Db 5380 TTTTCAAGGAAAGTTTCAAGGAGAGAGGCGAAGTTTTCAAAACATTTGTTTCAGAGAA 5439  
 Oy 2494 GGGAGCATTAAGTTTACAGCTTACAGAGAGCTTACACAAATATCTGTGTGAGAGAGCA 2553  
 Db 5440 GGGAGCATTAAGTTTACAGCTTACAGAGAGCTTACACAAATATCTGTGTGAGAGAGCA 5499  
 Oy 2554 GCATTTATCTATTTTATTTTATTAATAGTTTGTGCTTATCTTAAATGAATTTAAT 2613  
 Db 5500 GCATTTATCTATTTTATTTTATTAATAGTTTGTGCTTATCTTAAATGAATTTAAT 5559  
 Oy 2614 GTCACAAACTGTACACAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2666  
 Db 5560 GTCACAAACTGTACACAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5612

RESULT 14  
 AK090948  
 LOCUS  
 DEFINITION Homo sapiens cDNA FLJ33629 fis, clone BRAM2022454, highly similar  
 to Homo sapiens cell cycle checkpoint protein CHFR mRNA.  
 AK090948  
 ACCESSION AK090948.1 GI:21749208  
 VERSION oligo capping; fis (full insert sequence).  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1  
 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,  
 Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,  
 Sekine, M., Obase, M., Nishi, T., Shibahara, T., Tanaka, T.,  
 Iehli, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,  
 Nagahara, K., Murakami, K., Yaeuda, T., Iwayanagi, T., Wagatsuma, M.,  
 Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,  
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,  
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsura, N., Sato, K.,  
 Tanikawa, M., Yamazaki, K., Nishimura, K., Iehabashi, T., Yamauchi, H.,  
 Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,  
 Hiraoaka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,  
 Yoshida, M., Hottu, T., Kusano, J., Kanehori, K., Takahashi, F., Hara, R.,  
 Takeuchi, K., Arita, M., Imose, N., Muesashino, K., Yuki, H., Oshima, A.,  
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,  
 Shiohata, N., Sano, S., Moriya, S., Momiya, H., Satoh, N., Takami, S.,  
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,  
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,  
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,  
 Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,

Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobayashi, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Send, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Saitoh, T., Shitai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuko, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs  
Nat. Genet. 36 (1), 40-45 (2004)  
14702039

2  
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuko, Y., Nagai, K. and Isegai, T. NEDO human cDNA sequencing project  
Unpublished

3 (bases 1 to 2257)  
Isegai, T. and Yamamoto, J. Direct Submission  
Submitted (04-JUL-2002) Takao Isegai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genome@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.  
Location/Qualifiers

1. 2257  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRAMT2022454"  
/issue\_type="amygdala"  
/clone\_id="BRAMT2"  
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 35.9%; Score 961.8; DB 9; Length 2257;  
Best Local Similarity 99.8%; Pred. No. 1.7e-211;  
Matches 963; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1702 GAGCTCAACCTGGGAGCAAGTGTCTGAGCGGGGCTGTAACAACAAGCTACGAGTGA 1761  
669 GAGCTCAACCTGGGAGCAAGTGTCTGAGCGGGGCTGTAACAACAAGCTACGAGTGA 728  
1763 GACATCTGAGAAATTACTGTGGCAACGAGGTTTGACATGCAAAAAAAGCTTGAACCGAG 1821  
729 GACATCTGAGAAATTACTGTGGCAACGAGGTTTGACATGCAAAAAAAGCTTGAACCGAG 788  
1822 AACCTGTGGCTCTCCAGCGGGGAGTGTCTGTCTGTGATTAACAAGTCAACGGGAGAC 1881  
789 AACCTGTGGCTCTCCAGCGGGGAGTGTCTGTCTGTGATTAACAAGTCAACGGGAGAC 848  
1882 ACCGTTCTGTGTTACTGTGTGGCCGCGCAGCTTCGAGAGCTGACCTATGAGTGGG 1941  
849 ACCGTTCTGTGTTACTGTGTGGCCGCGCAGCTTCGAGAGCTGACCTATGAGTGGG 908  
1942 CAGAAATTCTGTCTCCAGAGTGGCCAGTGGCCGTAACAATCCGCTCTGTAACCTG 2001  
909 CAGAAATTCTGTCTCCAGAGTGGCCAGTGGCCGTAACAATCCGCTCTGTAACCTG 968

2002 GGCCTGAATCGCCGACATGAGTGAAGTCAACACGCGCATGAATTAATGATATCTGT 2061  
969 GGCCTGAATCGCCGACATGAGTGAAGTCAACACGCGCATGAATTAATGATATCTGT 1028  
2062 GAACAGCAACAGTTCAAAAACCTAGCATTCAGAGGCCCTGACAGCTTTCAGCACTGAG 2121  
1029 GAACAGCAACAGTTCAAAAACCTAGCATTCAGAGGCCCTGACAGCTTTCAGCACTGAG 1088  
2122 GTGAAGAGAGCGTGTGTTTAAATACAGAGACAGACAGCTCAAGGTGTTTCAACAGCCC 2181  
1089 GTGAAGAGAGCGTGTGTTTAAATACAGAGACAGACAGCTCAAGGTGTTTCAACAGCCC 1148  
2182 CTGAGGAAAGGAGACCGAGGCTTCGACAGAGGCTCTGGGGGAGCTCTCTGTGAGCTT 2241  
1149 CTGAGGAAAGGAGACCGAGGCTTCGACAGAGGCTCTGGGGGAGCTCTCTGTGAGCTT 1208  
2242 TTACCTCTGTGAGTGAAGACCTCCCGAGAGCCCGGGGGCCGACGCCCTCTGTGAG 2301  
1209 TTACCTCTGTGAGTGAAGACCTCCCGAGAGCCCGGGGGCCGACGCCCTCTGTGAG 1268  
2302 AGCGCTGGGACAGGCTCTGTGTGGCATGACAGACAGACGAAAGCTTCTGTAAATGC 2361  
1269 AGCGCTGGGACAGGCTCTGTGTGGCATGACAGACAGACGAAAGCTTCTGTAAATGC 1328  
2362 GGCCTGCCCGCCGACAGAGGGGAGTGTGCTCTTTGTACATTTTCGAAACCTACAGTTA 2421  
1329 GGCCTGCCCGCCGACAGAGGGGAGTGTGCTCTTTGTACATTTTCGAAACCTACAGTTA 1388  
2422 AACAGAGCTGTGTTTCAAGAAAGTTTCAAGGAGAAAGGAGGCAAGTTATCAAAACATT 2481  
1389 AACAGAGCTGTGTTTCAAGAAAGTTTCAAGGAGAAAGGAGGCAAGTTATCAAAACATT 1448  
2482 GTTTCAGAGAAAGAGACATAAGTTTACAGCTTACAGACGTACCAATATCTGTGCT 2541  
1449 GTTTCAGAGAAAGAGACATAAGTTTACAGCTTACAGACGTACCAATATCTGTGCT 1508  
2542 GGGAAACACAGCAATTTTATCTATTTTATTTATATAGTGTGTATCTTTCTTA 2601  
1509 GGGAAACACAGCAATTTTATCTATTTTATTTTATATAGTGTGTATCTTTCTTA 1568  
2602 TAAAGTTTAAATGTCAACAACTGTGACCAAAATTAATTAATTAATTAATTAATGAC 2661  
1569 TAAAGTTTAAATGTCAACAACTGTGACCAAAATTAATTAATTAATTAATTAATGAC 1628  
2662 AAAA 2666  
1629 TAAA 1633

RESULT 15  
AX713476 2186 bp DNA linear PAT 15-APR-2003  
LOCUS AX713476  
DEFINITION Sequence 160 from Patent EP1293569.  
ACCESSION AX713476  
VERSION AX713476.1 GI:29888332  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homosapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuko, Y. Full-length cDNAs  
Patent: EP 1293569-A 160 19-MAR-2003;  
Helix Research Institute (JP); Research Association for Biotechnology (JP)  
Location/Qualifiers  
1. 2186  
/organism="Homo sapiens"

FEATURES  
source

## ORIGIN

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Query Match 32.7%; Score 876.2; DB 6; Length 2186;  
Best Local Similarity 99.4%; Pred. No. 1.1e-191;  
Matches 890; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1772 AGAATTACCTGGCAACAGAGTTTGACATGAAAAACATGTTGACCGAGAGCTGTGG 1831  
DB 671 AGAATTACCTGGCAACAGAGTTTGACATGAAAAACATGTTGACCGAGAGCTGTGG 730  
QY 1832 CTCTCCAGGGGGAGGTTTCTGCTGTGATTAACAGAGTCAAGGAGACACCGTTCTGT 1891  
DB 731 CTCTCCAGGGGGAGGTTTCTGCTGTGATTAACAGAGTCAAGGAGACACCGTTCTGT 790  
QY 1892 GTTACTGCTGTGGCTGCGCAGGCTCCGTGAGCTGACCTATCATGATCGGAGAAATTC 1951  
DB 791 GTTACTGCTGTGGCTGCGCAGGCTCCGTGAGCTGACCTATCATGATCGGAGAAATTC 850  
QY 1952 CTGCTTCCGAGTTGCCAGTGGCCGTAACATCCCGTCTGACTGTGAGGCGGTAACT 2011  
DB 851 CTGCTTCCGAGTTGCCAGTGGCCGTAACATCCCGTCTGACTGTGAGGCGGTAACT 910  
QY 2012 GCCGCACTAGGTAAGTCAACAGCCATGAAATTCATATCTGTGAACAGACAA 2071  
DB 911 GCCGCACTAGGTAAGTCAACAGCCATGAAATTCATATCTGTGAACAGACAA 970  
QY 2072 GGTTCAAAAATTAAGCATCCAGAGGCTTCAAGGCTTTCAAGGCTTTCAAGGCTTTCA 2131  
DB 971 GGTTCAAAAATTAAGCATCCAGAGGCTTCAAGGCTTTCAAGGCTTTCAAGGCTTTCA 1030  
QY 2132 CGTGTGTTTAAATACAGACAGACAGTCAAGGCTTTCAAGGCTTTCAAGGCTTTCA 2191  
DB 1031 CGTGTGTTTAAATACAGACAGACAGTCAAGGCTTTCAAGGCTTTCAAGGCTTTCA 1090  
QY 2192 GGACGAGGGGTCTCCGACAGTGTCTGAGGGTGAAGCTTCTGTGAGGCTTTTACCTCT 2251  
DB 1091 GGACGAGGGGTCTCCGACAGTGTCTGAGGGTGAAGCTTCTGTGAGGCTTTTACCTCT 1148  
QY 2252 GAGTGAACCTCCCGAGAGCCCGGGGGCCGAGCCGCTCTGTGAGCGCTGGGC 2311  
DB 1149 GAGTGAACCTCCCGAGAGCCCGGGGGCCGAGCCGCTCTGTGAGCGCTGGGC 1208  
QY 2312 AGGGCTCGTGTGGCATCAGCAGACAGACGAAAGCTTTCTGTAAATGCGGCGTCCG 2371  
DB 1209 AGGGCTCGTGTGGCATCAGCAGACAGACGAAAGCTTTCTGTAAATGCGGCGTCCG 1268  
QY 2372 CCGAGAGGGGAGTTTGTCTTTTGTACATTTCCGAAACTACAGTTAAAGCAGAGTC 2431  
DB 1269 CCGAGAGGGGAGTTTGTCTTTTGTACATTTCCGAAACTACAGTTAAAGCAGAGTC 1328  
QY 2432 TGTGTTCAAGAAAAGTTTCAGGGAGAAAGGCGAAGTTTATCAAAAACATGTTTCAGG 2491  
DB 1329 TGTGTTCAAGAAAAGTTTCAGGGAGAAAGGCGAAGTTTATCAAAAACATGTTTCAGG 1388  
QY 2492 AAGGAGCATTAAGTTTACAGCCTACAGACGTACACAATATCTGTCTGTGAAAAACCA 2551  
DB 1389 AAGGAGCATTAAGTTTACAGCCTACAGACGTACACAATATCTGTCTGTGAAAAACCA 1448  
QY 2552 CAGCATTTTATCTATTTTATTTTAAATAGTTTGTGCTTATCTTAAATAGATTAA 2611  
DB 1449 CAGCATTTTATCTATTTTATTTTAAATAGTTTGTGCTTATCTTAAATAGATTAA 1508  
QY 2612 ATGTACAAACTGTACACAATAATATATTAATTTTACAAATGACAAAA 2666  
DB 1509 ATGTACAAACTGTACACAATAATATATTAATTTTACAAATGACAAAA 1563

Search completed: January 15, 2005, 04:02:36  
Job time : 11105 secs



XX The present sequence is that of cDNA encoding the human mitotic  
 CC checkpoint protein Chfr (see AAB20219) having a forkhead associated  
 CC domain (FHA) and a ring finger domain. The protein is required for  
 CC regulation of the transition of cells from prophase to metaphase during  
 CC mitosis. It has ubiquitin-protein ligase activity. The Chfr checkpoint  
 CC was evident in primary human cells, but was inactivated in 4 of 8 human  
 CC cancer cell lines. In U2OS cells, a mutation was identified, which caused  
 CC a Val to Met amino acid substitution in the highly conserved C-terminal  
 CC Cys-rich region of the Chfr protein. In the absence of the Chfr  
 CC checkpoint, cells subjected to mitotic stress condensed their chromosomes  
 CC despite failing to separate their chromosomes. Chfr may monitor  
 CC centrosome separation. Inactivation of the chfr gene in human cancer is  
 CC theorized to underlie the increased sensitivity of cancer cells to  
 CC antimetabolic drugs. Nucleic acids comprising the present sequence, or  
 CC sequences encoding at least amino acids 31-103, 303-346, 476-641 (or  
 CC their antisense sequences) are claimed. The Chfr cDNA was isolated from  
 CC an expressed sequence tag database sequence for cDNAs with FHA motifs.  
 CC Claimed methods of determining the tumorigenic potential of a cell  
 CC comprise examining the cell for the presence of Chfr nucleic acid  
 CC (absence or mutation indicating predisposition to tumorigenesis upon  
 CC exposure to mitotic stress). Also claimed is a diagnostic agent, e.g. an  
 CC antisense fragment of the present sequence, that binds to the Chfr  
 CC nucleic acid, and a diagnostic kit for detecting tumorigenic potential  
 CC of a cell. A composition which inhibits the biological activity of Chfr  
 CC may comprise a ligand selected from an antibody or its fragment. The Chfr  
 CC inhibitor is used in a claimed method of rearing the growth of a cancer  
 CC cells. Chfr polypeptides are useful in screening for drugs which can  
 CC inhibit the activity of Chfr in a cancer cell, rendering the cell more  
 CC sensitive to additional antitumour therapies

XX Sequence 2679 BP; 650 A; 715 C; 764 G; 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 2679; DB 5; Length 2679;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAAATTCGGACACGAGCCGCAATGCTTTGACAGCGCGCGCGCGAGTTCGG 60  
 DB 1 AAGAAATTCGGACACGAGCCGCAATGCTTTGACAGCGCGCGCGCGAGTTCGG 60  
 QY 61 GTTCG 120  
 DB 61 GTTCG 120  
 QY 121 CCG 180  
 DB 121 CCG 180  
 QY 181 GTTCG 240  
 DB 181 GTTCG 240  
 QY 241 CCCAGCAATTAACCTGCTCTGAGATCACTGTAATTTGATGATGATGATGAT 300  
 DB 241 CCCAGCAATTAACCTGCTCTGAGATCACTGTAATTTGATGATGATGATGAT 300  
 QY 301 CAGGTGACACTGGAAGATCCAGCACCAAGTGAACAGTATTAACAAGCTGAAG 360  
 DB 301 CAGGTGACACTGGAAGATCCAGCACCAAGTGAACAGTATTAACAAGCTGAAG 360  
 QY 361 AAGAGACGACATGCGCTTTACAGCTGGGATGTCATCTACTTGGTGTACAGGA 420  
 DB 361 AAGAGACGACATGCGCTTTACAGCTGGGATGTCATCTACTTGGTGTACAGGA 420  
 QY 421 GAACCGGAACAACAGTGGCATCTCTATGATCTTTAAGTAAAGCAAGGATGACA 480  
 DB 421 GAACCGGAACAACAGTGGCATCTCTATGATCTTTAAGTAAAGCAAGGATGACA 480  
 QY 481 CAAGATCTTTGAAGCTTACAGAGAAATGTTTCATGGGACCAAGATACCTCAG 540  
 DB 481 CAAGATCTTTGAAGCTTACAGAGAAATGTTTCATGGGACCAAGATACCTCAG 540

QY 541 CGAGGTGACAGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 DB 541 CGAGGTGACAGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600  
 QY 601 TGCCTTTGAGAAACCAACAGCCATCAACATGACAGTCAACCTTCCCGACAGCC 660  
 DB 601 TGCCTTTGAGAAACCAACAGCCATCAACATGACAGTCAACCTTCCCGACAGCC 660  
 QY 661 TCTTCCAGAGACCTTCTCTGACAGGCGAGAGCGCTTCTCAGTGTGTGGTGG 720  
 DB 661 TCTTCCAGAGACCTTCTCTGACAGGCGAGAGCGCTTCTCAGTGTGTGGTGG 720  
 QY 721 GGTGACATCTCCCTTAAGAGAGTGTCTCTGTGTGCAATGATGAGTCTCCAG 780  
 DB 721 GGTGACATCTCCCTTAAGAGAGTGTCTCTGTGTGCAATGATGAGTCTCCAG 780  
 QY 781 GCTCAGCTCTCCAGACAGAAACAGCGCTTCTTGTGTGTGGAACCCAGAGAT 840  
 DB 781 GCTCAGCTCTCCAGACAGAAACAGCGCTTCTTGTGTGTGGAACCCAGAGAT 840  
 QY 841 GAGATTTGAGCCCGTGAAGAGAAATGAGAGAGATGAGAGACCTTGAACCG 900  
 DB 841 GAGATTTGAGCCCGTGAAGAGAAATGAGAGAGATGAGAGACCTTGAACCG 900  
 QY 901 CAGTTGTGTGTGCAACACCGCGTGAAGATGCGCAACCGTCCACAGAGATG 960  
 DB 901 CAGTTGTGTGTGCAACACCGCGTGAAGATGCGCAACCGTCCACAGAGATG 960  
 QY 961 GCGCGTGGAGAGCAGACAGAAATGAGAGAGAGCGTGAATCATCATATCTG 1020  
 DB 961 GCGCGTGGAGAGCAGACAGAAATGAGAGAGAGCGTGAATCATCATATCTG 1020  
 QY 1021 CTGACACATCTGCTGAGTTCGACGCTTCGATGACACGTTTCGCGCGCTT 1080  
 DB 1021 CTGACACATCTGCTGAGTTCGACGCTTCGATGACACGTTTCGCGCGCTT 1080  
 QY 1081 GGTGTGATGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 DB 1081 GGTGTGATGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 QY 1141 AAAAACAACATCTCTCAACAACTCTGAGAGATCATCTCATCCAGATCCAGA 1200  
 DB 1141 AAAAACAACATCTCTCAACAACTCTGAGAGATCATCTCATCCAGATCCAGA 1200  
 QY 1201 CCGAGTGAAGAGATGTCACAAAGTATGATGATGATGATGATGATGATGAT 1260  
 DB 1201 CCGAGTGAAGAGATGTCACAAAGTATGATGATGATGATGATGATGATGAT 1260  
 QY 1261 CAGCCCAAGTCAAGCGGCTTTTCTGATGAAAGAGAGGATTCAGAGACCT 1320  
 DB 1261 CAGCCCAAGTCAAGCGGCTTTTCTGATGAAAGAGAGGATTCAGAGACCT 1320  
 QY 1321 CTGTGACAGCTTGAACAGTGTCTCTCAGACATTAAGCCAGCATTCAGTGT 1380  
 DB 1321 CTGTGACAGCTTGAACAGTGTCTCTCAGACATTAAGCCAGCATTCAGTGT 1380  
 QY 1381 TGTCTTGAATGACAGAGCGAGCGCGAGCTTCCCATCTGCGCAGCACTCC 1440  
 DB 1381 TGTCTTGAATGACAGAGCGAGCGCGAGCTTCCCATCTGCGCAGCACTCC 1440  
 QY 1441 CAGAGAGCCCAACAGGCGCTGAGAGATGACACCTTCCAGTGTCTGACAGCA 1500  
 DB 1441 CAGAGAGCCCAACAGGCGCTGAGAGATGACACCTTCCAGTGTCTGACAGCA 1500  
 QY 1501 GTCCAGAGATTAACGTGTGCTCTGACAGAGACCAAGCCCTGTGACCTGT 1560  
 DB 1501 GTCCAGAGATTAACGTGTGCTCTGACAGAGACCAAGCCCTGTGACCTGT 1560  
 QY 1561 CCGATGCCCCAGCCGAGAGCGAGAGCGCGAGACCGCGGTGTGCGCCCTC 1620  
 DB 1561 CCGATGCCCCAGCCGAGAGCGAGAGCGCGAGACCGCGGTGTGCGCCCTC 1620  
 QY 1621 GCGGTGTGCTGACAGCTTTCTGTGCACTGTGAGTGGGCTGCAACCGGAG 1680



```

Db      1621 GGGGTCTGCTGAGCCTTTCTGTCACCTGTACTGGGGCTGACCCGGACCCGGCTCTAC 1680
Qy      1681 GGTGCTGCTGGCCCCCTTTTGTGAGCTCAACCTGGGTGACAGAGTGTGTGACCGCCGCTG 1740
Db      1681 GGTGCTGCTGGCCCCCTTTTGTGAGCTCAACCTGGGTGACAGAGTGTGTGACCGCCGCTG 1740
Qy      1741 AACCAACAGAGCTACGAGTCAAGCATCTCTGAAGATTACTGGCAACAGAGTTTGACA 1800
Db      1741 AACCAACAGAGCTACGAGTCAAGCATCTCTGAAGATTACTGGCAACAGAGTTTGACA 1800
Qy      1801 TGGAAAAACATGTGTGACCGAGACCTCTGCTCTCAGCGGGAGTGTTCCTGCTGCT 1860
Db      1801 TGGAAAAACATGTGTGACCGAGACCTCTGCTCTCAGCGGGAGTGTTCCTGCTGCT 1860
Qy      1861 GATTACAGAGTCACGGGAGACACCGTCTGTGTGATCTGCTGTGGCTGGCAGCTCCCT 1920
Db      1861 GATTACAGAGTCACGGGAGACACCGTCTGTGTGATCTGCTGTGGCTGGCAGCTCCCT 1920
Qy      1921 GAGCTGACCTATCAGATGATGGGAGAACATTCCTGCTCCGAGTTGCAGTGGCCGTTACA 1980
Db      1921 GAGCTGACCTATCAGATGATGGGAGAACATTCCTGCTCCGAGTTGCAGTGGCCGTTACA 1980
Qy      1981 TCCCGTCTGACTGCTACTGAGGCGCGTAACTGCGGACTCAGGTGAAGCTCAACAGCC 2040
Db      1981 TCCCGTCTGACTGCTACTGAGGCGCGTAACTGCGGACTCAGGTGAAGCTCAACAGCC 2040
Qy      2041 ATGAATTCATATCATATCTGTGTGAACAGACAGAGTTCAAAAATCAATCCAGAGCCCT 2100
Db      2041 ATGAATTCATATCATATCTGTGTGAACAGACAGAGTTCAAAAATCAATCCAGAGCCCT 2100
Qy      2101 GAGCAGCTTTCAGCAGTGAAGAGAGCGCTTTTAAATTCAGAGCAACAGCAG 2160
Db      2101 GAGCAGCTTTCAGCAGTGAAGAGAGCGCTTTTAAATTCAGAGCAACAGCAG 2160
Qy      2161 TCAAGGTGTTTTCACAGCCCGCTGAGGGAAGGAGCCAGGGTCTCCGAGAGTGTCTG 2220
Db      2161 TCAAGGTGTTTTCACAGCCCGCTGAGGGAAGGAGCCAGGGTCTCCGAGAGTGTCTG 2220
Qy      2221 GGTGACTCTTCTGTGAGAGCTTTTAACTCTGTGATGAGACCTCCCAAGCCCGGGG 2280
Db      2221 GGTGACTCTTCTGTGAGAGCTTTTAACTCTGTGATGAGACCTCCCAAGCCCGGGG 2280
Qy      2281 CGGAGCGCGCCCTCTGTGAGAGCGTGGGCGGGCTCGGTGATGAGATCAGACAGAGA 2340
Db      2281 CGGAGCGCGCCCTCTGTGAGAGCGTGGGCGGGCTCGGTGATGAGATCAGACAGAGA 2340
Qy      2341 CGAAGCCTTCTGTAAATGCGGCGCTCCCGCGAGAGGGGCACTTCTCTTTGTAC 2400
Db      2341 CGAAGCCTTCTGTAAATGCGGCGCTCCCGCGAGAGGGGCACTTCTCTTTGTAC 2400
Qy      2401 ATTTTCGAAACCTACAGTTAAAGCAGAGTCTGTTTTCAGAAAAGTTTCAAGAGAG 2460
Db      2401 ATTTTCGAAACCTACAGTTAAAGCAGAGTCTGTTTTCAGAAAAGTTTCAAGAGAG 2460
Qy      2461 GCGAAGTTTATCAAAAAATGTTTTCAGAGAAAGGAGATTAAGCTTAAGAG 2520
Db      2461 GCGAAGTTTATCAAAAAATGTTTTCAGAGAAAGGAGATTAAGCTTAAGAG 2520
Qy      2521 CGTACCAATATCTCTGCTGCGGAAACACAGCATTTTATATTTTATTTAA 2580
Db      2521 CGTACCAATATCTCTGCTGCGGAAACACAGCATTTTATATTTTATTTAA 2580
Qy      2581 GGTGTTGCTTATCTTCTAATTAAGATTTAATGTCACAACTGTAGCAAAATTA 2640
Db      2581 GGTGTTGCTTATCTTCTAATTAAGATTTAATGTCACAACTGTAGCAAAATTA 2640
Qy      2641 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2699
Db      2641 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2699
Qy      2699 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2699
Db      2699 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2699

```

```

AAH14542
ID      AAH14542 standard; cDNA; 2639 BP.
XX
AC      AAH14542;
XX
DT      26-JUN-2001 (first entry)
XX
DE      Human cDNA sequence SEQ ID NO:12099.
XX
KW      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      EP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-00116126.
XX
PR      29-JUL-1999; 994P-00248036.
PR      27-ANG-1999; 993P-00300253.
PR      11-JAN-2000; 2000JP-00118776.
PR      02-MAY-2000; 2000JP-00183767.
PR      09-JUN-2000; 2000JP-00241899.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR      WPI; 2001-318749/34.
XX
PT      Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT      length cDNAs defined in the specification, and for the detection and/or
PT      diagnosis of the abnormality of the proteins encoded by the full-length
PT      cDNAs.
XX
PS      Claim 8; SEQ ID NO 12099; 2537pp + Sequence Listing; English.
XX
CC      The present invention describes primer sets for synthesizing 5602 full-
CC      length cDNAs defined in the specification. Where a primer set comprises:
CC      (a) an oligo-dT primer and an oligonucleotide complementary to the
CC      complementary strand of a polynucleotide which comprises one of the 5602
CC      nucleotide sequences defined in the specification, where the
CC      oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC      of an oligonucleotide comprising a sequence complementary to the
CC      complementary strand of a polynucleotide which comprises a 5'-end
CC      sequence and an oligonucleotide comprising a sequence complementary to a
CC      polynucleotide which comprises a 3'-end sequence, where the
CC      oligonucleotide comprises at least 15 nucleotides and the combination of
CC      the 5'-end sequence/3'-end sequence is selected from those defined in the
CC      specification. The primer sets can be used in antisense therapy and in
CC      gene therapy. The primers are useful for synthesizing polynucleotides,
CC      particularly full-length cDNAs. The primers are also useful for the
CC      detection and/or diagnosis of the abnormality of the proteins encoded by
CC      the full-length cDNAs. The primers allow obtaining of the full-length
CC      cDNAs easily without any specialized methods. AAH0166 to AAH13628 and
CC      AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC      represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC      oligonucleotides, all of which are used in the exemplification of the
CC      present invention.
XX
SQ      Sequence 2639 BP; 620 A; 706 C; 760 G; 553 T; 0 U; 0 Other;
XX
Query Match      95.8%; Score 2566.4; DB 4; Length 2639;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2613; Conservative 0; Mismatches 1; Indels 36; Gaps 1;
Qy      17 GCCGCAATGCTCTTGTGACAGCGCGCGCGAGCGGTTCCGGGTTCCGGCGGCGGCGG 76
Db      5 GCCGCAATGCTCTTGTGACAGCGCGCGCGAGCGGTTCCGGGTTCCGGCGGCGGCGG 64
Qy      77 GGATGGAATCCGATGAGCGCGCCGAGGAAGCAAGATGCGCGCGCGGCAAGCCT 136

```

Db 65 GAGTGGAAATCCGAGTGGAGCGGCGCCGAGAAAGCAAGCAAGTCGCGCGCGCAAGCCTT 124  
Qy 137 GGGGACGAGTCCCTGCGCTGCGGCGCGAGAGAGGGCGAGCCGCAAGCTCTCTGAGGAAGC 196  
Db 125 GGGGAGGGGTCTCGTGGCTCTGGGCGCGAGAGAGGGCGAGCCGCAAGCTCTCTGAGGAAGC 184  
Qy 197 GGGAGTGAACCATCGGGCGGAGACGAGGTTGCGACCTTTCTTTCCGAGCAATTAACCTGG 256  
Db 185 GGGAGTGAACCATCGGGCGGAGACGAGTGGCGACCTTTCTTTCCGAGCAATTAACCTGG 244  
Qy 257 TCTCTGAGATCATCTGATGAATTTGTAGTGAATGAATAATCAGGTGAGGATGAGCACTGGAA 316  
Db 245 TCTCTGAGATCATCTGATGAATTTGTAGTGAATGAATAATCAGGTGAGGATGAGCACTGGAA 304  
Qy 317 ATACGACGACAGTGGAGAGTGAATTAACAAGCTGAAGGTGTTAAGAAGCAGACATGCC 376  
Db 305 ATACGACGACAGTGGAGAGTGAATTAACAAGCTGAAGGTGTTAAGAAGCAGACATGCC 364  
Qy 377 CTTTACAGACTGGGAGTGTCACTTCTGCTGTAACGAAAGATGAACCGGAACACAACG 436  
Db 365 CTTTACAGACTGGGAGTGTCACTTCTGCTGTAACGAAAGATGAACCGGAACACAACG 424  
Qy 437 TGGCATACCTCATGAAATCTTTAAGTGAAGCAAGGAGCATGACAAAGAACTCTTGAAG 496  
Db 425 TGGCATACCTCATGAAATCTTTAAGTGAAGCAAGGAGCATGACAAAGAACTCTTGAAG 480  
Qy 497 CTAAAGAGAAATGTGTTCATGGGACCAAGATACCTCAGGTGAGGTGAGGCGAG 556  
Db 481 -----GATACCTCAGGTGAGGTGAGGCGAG 508  
Qy 557 GGGGCGATCCCGGGGTCCCTCCGTGCGTGGCCCGGACATGAGTGTCTTTGAGAACAC 616  
Db 509 GGGGCGATCCCGGGGTCCCTCCGTGCGTGGCCCGGACATGAGTGTCTTTGAGAACAC 568  
Qy 617 AGCCATCAACATGACGTGAGACCTTTCCCAAGCCCTGGGCTCTTCAAGAGGCTT 676  
Db 569 AGCCATCAACATGAGTGAACCTTTCCCAAGCCCTGGGCTCTTCAAGAGGCTT 628  
Qy 677 CTCCTGAGAGGCGAGAGCGTTCCTCCAGTTGTGGGTCTGGGGGTGGTGGCATCTCCCTA 736  
Db 629 CTCCTGAGAGGCGAGAGCGTTCCTCCAGTTGTGGGTCTGGGGGTGGTGGCATCTCCCTA 688  
Qy 737 AAGGAATGATCCCTCTGTGGAAGTGAAGTCTCCAGCTTTGCTTCAAGCTCTCCAG 796  
Db 689 AAGGAATGATCCCTCTGTGGAAGTGAAGTCTCCAGCTTTGCTTCAAGCTCTCCAG 748  
Qy 797 ACAGAAAGCTGCTCTTTCTGTGTTGAACCCAGGATCAGAGAGATTTGAGAGCCG 856  
Db 749 ACAGAAAGCTGCTCTTTCTGTGTTGAACCCAGGATCAGAGAGATTTGAGAGCCG 808  
Qy 857 TGAAGAGAAATGAGAGAGATGGGAGCTTTGACCTGAACGGGACGTTGTGTGCGAC 916  
Db 809 TGAAGAGAAATGAGAGAGATGGGAGCTTTGACCTGAACGGGACGTTGTGTGCGAC 868  
Qy 917 AACCGGTGAAGATGCCCAACCGTCCAGAGAGATCAGAGACCGGCTGGGAAGCCAG 976  
Db 869 AACCGGTGAAGATGCCCAACCGTCCAGAGAGATCAGAGACCGGCTGGGAAGCCAG 928  
Qy 977 ACAGATGAGAGAGAGCTGATCATCTGATGAGGACCTTGTGACAGACTGCTGA 1036  
Db 929 ACAGATGAGAGAGAGCTGATCATCTGATGAGGACCTTGTGACAGACTGCTGA 988  
Qy 1037 GTTTCAGGCTGATGACAGCTTCTGCGCGGCTTGTCTACTCGGCGTGAATGAGGCT 1096  
Db 989 GTTTCAGGCTGATGACAGCTTCTGCGCGGCTTGTCTACTCGGCGTGAATGAGGCT 1048  
Qy 1097 GGTCCCTGTGTCTACTGCGCGCTGCTCCGTTGGAAGGATCTGTGAATAACCATCTCA 1156  
Db 1049 GGTCCCTGTGTCTACTGCGCGCTGCTCCGTTGGAAGGATCTGTGAATAACCATCTCA 1108  
Qy 1157 ACAACCTGTGAGACATACCTCATCAGACATCAGAGAGAGTGCAGTGAAGAGATG 1216

Db 1109 ACAACCTGTGAGACATACCTCATCAGACATCAGAGAGAGTGCAGTGAAGAGATG 1168  
Qy 1217 TGAAGATGATGATGCCAGATTAATAATCACTCAAGACATGCTGAGCCCAAGTCAAGC 1276  
Db 1169 TGAAGATGATGATGCCAGATTAATAATCACTCAAGACATGCTGAGCCCAAGTCAAGC 1228  
Qy 1277 GGTCTTTTCTGATGAAGAGAGGATTCAGAGACCTGCTGAGAGTGTCAAGCTTGA 1336  
Db 1229 GGTCTTTTCTGATGAAGAGAGGATTCAGAGACCTGCTGAGAGTGTCAAGCTTGA 1288  
Qy 1337 GTGAGTCTCAGACATTAAGCAGACATACGTCGTGCGGAGAGTGTCAAGTGAAGAA 1396  
Db 1289 GTGAGTCTCAGACATTAAGCAGACATACGTCGTGCGGAGAGTGTCTGAAGTGAAGAA 1348  
Qy 1397 GAGAGCGCGCAGGCTCCCATGCGCAGACACCGAGGCGAGGCGAGAGCCCAAGC 1456  
Db 1349 GAGAGCGCGCAGGCTCCCATGCGCAGACACCGAGGCGAGGCGAGAGCCCAAGC 1408  
Qy 1457 CCTGGGGGATGACCTCTCAGTCCGTCAGCTGAGCAGACAGTCCAGATTAAGTGT 1516  
Db 1409 CCTGGGGGATGACCTCTCAGTCCGTCAGCTGAGCAGACAGTCCAGATTAAGTGT 1468  
Qy 1517 GCCCTGCAAGAAAGCCAGCCGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1576  
Db 1469 GCCCTGCAAGAAAGCCAGCCGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1528  
Qy 1577 GAGCGAGCGGAGACAGGACCCGCGTGTGCGCTCTCAGACATGTGCGGTCTGCTGAGC 1636  
Db 1529 GAGCGAGCGGAGACAGGACCCGCGTGTGCGCTCTCAGACATGTGCGGTCTGCTGAGC 1588  
Qy 1637 CTTTCTGACACTGTACTGCGGGCTGACACCGGACCGGCTGTGAAGGCTGCTGCGCCGT 1696  
Db 1589 CTTTCTGACACTGTACTGCGGGCTGACACCGGACCGGCTGTGAAGGCTGCTGCGCCGT 1648  
Qy 1697 TTTGTGAGCTCAACCTGGGATGAACAAGTGTGGAACGGGCTGTGAACAACAAGCTAG 1756  
Db 1649 TTTGTGAGCTCAACCTGGGATGAACAAGTGTGGAACGGGCTGTGAACAACAAGCTAG 1708  
Qy 1757 AGTCAGACATCTGAAGAAATTAACCTGGCAACAGAGGTTTGAATGAAAAATGTTGA 1816  
Db 1709 AGTCAGACATCTGAAGAAATTAACCTGGCAACAGAGGTTTGAATGAAAAATGTTGA 1768  
Qy 1817 CCGAGAGCTGCTGAGCTTCCAGGGGAGAGTGTCTGCTGTGATTAAGAGTCAAGC 1876  
Db 1769 CCGAGAGCTGCTGAGCTTCCAGGGGAGAGTGTCTGCTGTGATTAAGAGTCAAGC 1828  
Qy 1877 GAGACACGTTCTGTGTTAAGCTGTGAGGCTGCGACGTTCCGTAAGCTGATCAGT 1936  
Db 1829 GAGACACGTTCTGTGTTAAGCTGTGAGGCTGCGACGTTCCGTAAGCTGATCAGT 1888  
Qy 1937 ATCGGACGAACATTCCTGCTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 1996  
Db 1889 ATCGGACGAACATTCCTGCTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 1948  
Qy 1997 ACTGGGCGCGTAATCGCGGACCTCAGGTGAAGTCAACCGCATGAATTAATCATA 2056  
Db 1949 ACTGGGCGCGTAATCGCGGACCTCAGGTGAAGTCAACCGCATGAATTAATCATA 2008  
Qy 2057 TCTGTGAACAGACAGGTTCAAAAATTAAGCATCAGAGGCTTGAAGCTTCAAGC 2116  
Db 2009 TCTGTGAACAGACAGGTTCAAAAATTAAGCATCAGAGGCTTGAAGCTTCAAGC 2068  
Qy 2117 TGAAGTGAAGAGAGCTGTTTAAATTAACAGACAGACAGTCAAGGTTTTCACA 2176  
Db 2069 TGAAGTGAAGAGAGCTGTTTAAATTAAGCATCAGAGGCTTGAAGGTTTTCACA 2128  
Qy 2177 GCCCCTGAGGGAAGGAGACGAGGCTCCGACAGAGGCTCTGGGGGTGAGCTCTTCTGAG 2236  
Db 2129 GCCCCTGAGGGAAGGAGACGAGGCTCCGACAGAGGCTCTGGGGGTGAGCTCTTCTGAG 2188  
Qy 2237 AGCTTTTAACTCTGAGTGAAGCCTCCCAAGAGCCCGAGGAGCCCAAGCCGCTCC 2296  
Db 2189 AGCTTTTAACTCTGAGTGAAGCCTCCCAAGAGCCCGAGGAGCCCAAGCCGCTCC 2248

Qy	2287	GGGTAGGGCTGGGCGAGGGCTCGTGGTGGCAATCAGCAGACAGAGCAAGACCTTTCTGTAA	2356
Db	2249	TGGTAGGCGCTGGGCGAGGGCTCTGTGGTGCATCAGCAGACAGAGCAAGACCTTTCTGTAA	2308
Qy	2357	CATGCGGCGGTCCCGCCGAGAGGGCAGTTTGTCTCTTTGTACATTTTCCGAACCTACA	2416
Db	2309	CATGCGGCGGTCCCGCCGAGAGGGCAGTTTGTCTTTTGTACATTTTCCGAACCTACA	2368
Qy	2417	GTTAAGCAGAGTCTGTTTTCAGGAAAAGTTTCAGGGAGAGGGCAAGTTTATCAAA	2476
Db	2369	GTTAAGCAGAGTCTGTTTTCAGGAAAAGTTTCAGGGAGAGGGCAAGTTTATCAAA	2428
Qy	2477	ACATTGTTTCAGGAGAGGGAGCAATATTTACAGCCCTACAGAGCTACAAATATCCTG	2536
Db	2429	ACATTGTTTCAGGAGAGGGAGCAATATTTACAGCCCTACAGAGCTACAAATATCCTG	2488
Qy	2537	CTGCTGGGAAAACCAAGCATTTTATCTAATTTTAAATTAATAGTTTGGTGTATCT	2596
Db	2489	CTGCTGGGAAAACCAAGCATTTTATCTAATTTTAAATTAATAGTTTGGTGTATCT	2548
Qy	2587	TCTAATAAGATTTAAATGTCACAAACTGTAGACAAATATATATTAATAATTACAA	2656
Db	2549	TCTAATAAGATTTAAATGTCACAAACTGTAGACAAATATATATTAATAATTACAA	2608
Qy	2657	TTGACAAAAA 2666	
Db	2609	TTGACTTAAA 2618	

CC	XX	AAH14556	standard; cDNA; 3138 BP.
CC	XX	AAH14556;	
CC	XX	26-JUN-2001	(first entry)
CC	XX	Human cDNA	sequence SEQ ID NO:12127.
CC	XX	Human;	primer; detection; diagnosis; antisense therapy; gene therapy; ss
CC	XX	Homo sapiens.	
CC	XX	EP1074617-A2.	
CC	XX	07-FEB-2001.	
CC	XX	28-JUL-2000;	2000BP-00116126.
CC	XX	29-JUL-1999;	99JP-00248036.
CC	XX	27-AUG-1999;	99JP-00300253.
CC	XX	11-JAN-2000;	2000JP-00118776.
CC	XX	02-MAY-2000;	2000JP-00183767.
CC	XX	09-JUN-2000;	2000JP-00241899.
CC	XX	(HELI-) HELIX	RES INST.
CC	XX	Oca T,	Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
CC	XX	Ishii S,	Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
CC	XX	WPI;	2001-318749/34.
CC	XX	Primer sets	for synthesizing polynucleotides, particularly the 5602 full-
CC	XX	length cDNAs	defined in the specification, and for the detection and/or
CC	XX	diagnosis	of the abnormality of the proteins encoded by the full-length
CC	XX	cDNAs.	
CC	XX	Claim 8;	SEQ ID NO 12127; 2537pp + Sequence Listing; English.
CC	XX	The present	invention describes primer sets for synthesizing 5602 full-
CC	XX	length cDNAs	defined in the specification. Where a primer set comprises:
CC	XX	(a) an oligo-dT	primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH016 to AAH1528 and  
CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
CC represent human acid sequences; and AAH16529 to AAH16532 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

XX  
XX Sequence 3138 BP; 738 A; 824 C; 873 G; 703 T; 0 U; 0 Other;

Query Match 84.4%; Score 2262.4; DB 4; Length 3138;  
Best Local Similarity 93.0%; Pred. No. 0;  
Matches 2454; Conservative 0; Mismatches 61; Indels 125; Gaps 3

Query Match	84.4%	Score 2262.4	DB 4	Length 3138
Best Local Similarity	93.0%	Pred. No. 0		
Matches 2454	Conservative	0	Mismatches 61	Indels 125
			Gaps	3
QY	27	CTCTTACAGCGCGCGCGCGCGCGCGGTTCCGGTTTCGGCGCGCGCGCGGAGTGAAT	86	
Db	1	CTCTTACAGCGCGCGCGCGCGCGCGGTTCCGGTTTCGGCGCGCGCGGAGTGAAT	60	
QY	87	CCCGATGAGCGGCCCCGAGAAAGGCAAGCAGTCGCCCCGCGCAAGCCCTGGGGACGGCT	146	
Db	61	CCCGATGAGCGGCCCCGAGAAAGGCAAGCAGTCGCCCCGCGCAAGCCCTGGGGACGGCT	120	
QY	147	CCGCGCTCTGGGCGCGGAGAGGGGGAGCCGCAAGTCTCTCTGAGAGAGCGGAGTGGAC	206	
Db	121	CCGCGCTCTGGGCGCGGAGAGGGGGAGCCGCAAGTCTCTCTGAGAGAGCGGAGTGGAC	180	
QY	207	CATCGGGCGGAGAGCAGAGTTTGGACCTTTCCTCCCGACAAATTAAGTCTCTGGAGA	266	
Db	181	CATCGGGCGGAGAGCAGAGTTTGGACCTTTCCTCCCGACAAATTAAGTCTCTGGAGA	240	
QY	267	TCACTGTAAATTGTAGTGGATGAAATAATCAGGTCAAGTGAACCTGAAAGTACAGAC	326	
Db	241	TCACTGTAAATTGTAGTGGATGAAATAATCAGGTCAAGTGAACCTGAAAGTACAGAC	300	
QY	327	CAATGGAACAGGATTTAAACAAGCTGAAGTTTAAAGAGAGACATGCCCTTACAGAC	386	
Db	301	CAATGGAACAGGATTTAAACAAGCTGAAGTTTAAAGAGAGACATGCCCTTACAGAC	360	
QY	387	TGGGAGATGCATCTACTTGTGTGACAGGAAGATGAACCGGAAACAACAGTGCATCACT	446	
Db	361	TGGGAGATGCATCTACTTGTGTGACAGGAAGATGAACCGGAAACAACAGTGCATCACT	420	
QY	447	CTATGAATCTTTAAGTAAAGCAAGGATGACACAGAATCTTTGAAGCTAACAGGA	506	
Db	421	CTATGAATCTTTAAGTAAAGCAAGGATGACACAGAATCTTTGAAGTGCCTTG	480	
QY	507	AAATGTGTTCCATGGGACCAAGAAATCACTCAGGTGCAGGTGCAGGGCGAGGGCCATCC	566	
Db	481	CTGTGTGTTCCAGGGCTGTGTAAACTCTCGGGAATCAAGT	520	
QY	567	CCGGGTCCCTCGTGTGCGCCGCGCACTCAGGTGTCCTTGAGGAACACAGCATCAAC	626	
Db	521	---GATCTTCCACCTTGGCTTCCCAAAGTATGTGATTACA-----	559	
QY	627	ATGACGTGAGACTTCTTCCCAAGCCTCGGCTCTTCCACGAGGCTTCTCTGCAAG	686	
Db	560	-----	559	
QY	687	GCAAGAGCTTCTCCATGTTTGGGTCTGGGGGTGTGTGACATCTCCCTTAAAGAAATGG	746	
Db	560	-----GGGTCTGGGGGTGTGTGACATCTCCCTTAAAGAAATGG	597	

[illegible]

QY	1827	CGTGGCTCTCCAGCGGGGAGTGTTCGTCTGTCTGATTACAGAGTCAACGGAGACACCGT	1888
Db	1678	CGTGGCTCTCCAGCGGGGAGTGTTCGTCTGTCTGATTACAGAGTCAACGGAGACACCGT	1737
QY	1887	TCGTGTTACTAGCTGTGTGGCTCTGCGCAGCTTCGGTAGCTGACCTATCAGTATCGGCAGAA	1946
Db	1738	TCGTGTTACTAGCTGTGTGGCTCTGCGCAGCTTCGGTAGCTGACCTATCAGTATCAGAGAA	1797
QY	1947	CATTCTCTGCTTCCGAGTTGCCAGTGTGCGCCGTAAATCCCTGCTTGACTGTGCTACTGTGGGCGG	2006
Db	1798	CATTCTCTCTTCCGAGTTGCCAGTGTGCGCCGTAAATCCCTGCTTGACTGTGCTACTGTGGGCGG	1857
QY	2007	TAACTGCGGCACTGAGGTGAAGAGCTCAACAGCCATGAGAAATTCATATCATCTGTGAACA	2066
Db	1858	TAACTGCGGCACTGAGGTGAAGAGCTCAACAGCCATGAGAAATTCATATCATCTGTGAACA	1917
QY	2067	GACAAAGTTCAAAAATTAAGCATCCAGAGGCCCTTGAAGAGCTTTCAAGCACTGAGAGTGA	2126
Db	1918	GACAAAGTTCAAAAATTAAGCATCCAGAGGCCCTTGAAGAGCTTTCAAGCACTGAGAGTGA	1977
QY	2127	GAGAGCGGTGTTTTAAATTAAGAGACAGACGCTCAAGGTGTTTTCAACGCCCCCTGAG	2186
Db	1978	GAGAGCGGTGTTTTAAATTAAGAGAGACAGACGCTCAAGGTGTTTTCAACGCCCCCTGAG	2037
QY	2187	GGAAGGAGCGCAGGGGTCTCCGACAGGTGCTCTGAGGTGACTCTTCTGTGAGACTTTTTTAC	2246
Db	2038	GGAAGGAGCGCAGGGGTCTCCGACAGGTGCTCTGAGGTGACTCTTCTGTGAGAGC--TTTAC	2095
QY	2247	CCTCTGAATGAGACCTCTCCCAAGCCCCGGGGGGCCGACGCGCGCTCTGTGTGAGCGC	2306
Db	2096	CCTCTGAATGAGACCTCTCCCAAGCCCCGGGGGGCCGACGCGCGCTCTGTGTGAGAGCGC	2155
QY	2307	TGGGCAAGGCTCTGTGTGTGACATCAGCAGCAGAGACGAAAGCCCTTCTGTAAATAGCGGCG	2366
Db	2156	TGGGCAAGGCTCTGTGTGTGACATCAGCAGCAGAGACGAAAGCCCTTCTGTAAATAGCGGCG	2215
QY	2367	TCCCGCCGAGAGGGGCAAGTTTGCTCTTTGTATCAATTTCCGAAACTACAGTTAAAGCAG	2426
Db	2216	TCCCGCCGAGAGGGGCAAGTTTGCTCTTTGTATCAATTTCCGAAACTACAGTTAAAGCAG	2275
QY	2427	AAGTCTGTTTTCAGGAAAAAGTTTCAAGGGAGAGGGGCAAGTTTACAAAAACATTTGTC	2486
Db	2276	AAGTCTGTTTTCAGGAAAAAGTTTCAAGGGAGAGGGGCAAGTTTACAAAAACATTTGTC	2335
QY	2487	AGAGAAAGGAGACATAAGTTTACAGCCTTACAGCAGCTACCAATATCTGTGCTGGAA	2546
Db	2336	AGAGAAAGGAGACATAAGTTTACAGCCTTACAGCAGCTACCAATATCTGTGCTGGAA	2395
QY	2547	AACCAAGCANTTATATCTATTTTATTTTAAATAGTTTGTGTGTATCTTTCTAAATAGA	2606
Db	2396	AACCAAGCANTTATATCTATTTTATTTTAAATAGTTTGTGTGTATCTTTCTAAATAGA	2455
QY	2607	TTTAAATGTCACAACTGTAGACCAATATATATATTTATTAATAATGACAAAA	2666
Db	2456	TTTAAATGTCACAACTGTAGACCAATATATATATTTATTAATAATGACAAAA	2515
RESULT 4			
ABN59646			
ID	ABN59646	standard; cDNA; 3181 BP.	
AC	ABN59646;		
XX	28-JUN-2002	(first entry)	
DT	XX		
DE	XX	Novel human coding sequence SEQ ID NO: 57.	
XX	XX		
KM	Human; anti-neurotic; vulnerary; anti-inflammatory; immunomodulator;		
KM	anti-infectivity; cerebroprotective; cytoskeletal; rheumatic; gene therapy;		
KM	neuroprotective; anti-parkinsonian; protein therapy; EST;		
XX	expressed sequence tag; gene; ss.		
XX	Homo sapiens.		
OS			

PX	MO20222660-A2.
XX	
PD	21-MAR-2002.
XX	
PF	10-SEP-2001; 2001WO-US026015.
XX	
PR	11-SEP-2000; 2000US-00659671.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang Y <sup>T</sup> , Liu C, Zhou F, Asundi V, Zhang J, Zhao QA, Ren F;
PI	Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX	
PT	WPI; 2002-292408/33.
DR	P-PsDB; ABB97233.
XX	
PS	Claim 1; SEQ ID NO 57; 509bp; English.
XX	
CC	The present invention provides the protein and coding sequences of 444
CC	novel human proteins. These were isolated from expressed sequences tags
CC	(ESTs). They can be used to stimulate cell growth to regulate
CC	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC	fertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC	and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC	rheumatoid arthritis, and to treat nervous system disorders e.g.
CC	Parkinson's disease. The present sequence is a coding sequence of the
XX	Invention
SQ	Sequence 3181 BP; 756 A; 830 C; 879 G; 716 T; 0 U; 0 Other;
Query March	83.8%; Score 2244.6; DB 6; Length 3181;
Best Local Similarity	92.3%; Pred. No. 0;
Matches 2447; Conservative	0; Mismatches 79; Indels 125; Gaps 3
OY	16 GGCGGCAATGTCTCTTGAACAGCGGGCGCGGCAGCGGTTCCGGATTTCGCGCGGGGCG 75
Db	21 GGCGGTAACCGGTCGGGAATTCGGGGTGACGATTTCGTTCCGGGTTTCGGCGGGGGCG 80
OY	76 GGAATGTGAATCCCATGTGAGCGGCCCGAAGAAGCAGCATGCGCGCGCGCGAGCCC 135
Db	81 GGAATGTGAATCCCATGTGAGCGGCCCGAAGAAGCAGCATGCGCGCGCGCGAGCCC 140
OY	136 TGGGGACGGCTCTCTGCTGGGGCCGGAGGAGGGCGAGCGGCAAGTCTCTTAGAAG 195
Db	141 TGGGGACGGCTCTCTGCTGGGGCCGGAGGAGGGCGAGCGGCAAGTCTCTTAGAAG 200
OY	196 CGGAGTGGACCATCGGGCGAGAGAGGGTTTCGACTTTCTTCCCCGAAATAATCTG 255
Db	201 CGGAGTGGACCATCGGGCGAGAGAGGGTTTCGACTTTCTTCCCCGAAATAATCTG 260
OY	256 GTCTCTGAGATCACTGTAGAAATTGTAGTGAATGAAAAATCAGTCAAGTGAAGTGA 315
Db	261 GTCTCTGAGATCACTGTAGAAATTGTAGTGAATGAAAAATCAGTCAAGTGAAGTGA 320
OY	316 GATACACAGCACAGTGGAAACAATGATTTAACAGCTGAAGGTTGTTAAAGAGACATGC 375
Db	321 GATACACAGCACAGTGGAAACAATGATTTAACAGCTGAAGGTTGTTAAAGAGACATGC 380
OY	376 CCTTTACAGACTGGGAGTGTCACTTACTGTTGTTACAGGAAGATGAACCGGACACAAC 435
Db	381 CCTTTACAGACTGGGAGTGTCACTTACTGTTGTTACAGGAAGATGAACCGGACACAAC 440
OY	436 GTGGCAATCTCTATGAATCTTTAAAGTAAAAGCAAGCATGACCAAGATCTTTGAA 495
Db	441 GTGGCAATCTCTATGAATCTTTAAAGTAAAAGCAAGCATGACCAAGATCTTTGAG 500
OY	496 GCTAACAGAGAAATGTGTTCATGTGGACCAGAAAGTACTCTCAGTGTGCAAGGGGCA 555

Db	501	ATGTCCTCTGCTGTGTGTCCTGAGGCTGGTCTTAAATCTCTGGGATCAAGT-----	551
Qy	556	GAGGCGCATCCCGGGGTCCCTCCGTGTGCGCCGCCACTCAGAGTGTCTTTGAGAACCA	615
Db	552	-----GATCCTCCCACTTGGCCTCCCAAGTATTTGATTTACA-----	590
Qy	616	CAGGCATCAACATCGACGTCAAGCTCTTCCCAAGACGCTCGGCTCTTCCACGAGCCT	675
Db	591	-----	590
Qy	676	TCTCTCGAGGGCAGAGGTTCTTCAGTTGTGGGTCTGGGGGTGTGGGCATCTCCCT	735
Db	591	-----GGGTCTGGGGGTGTGGGCATCTCCCT	617
Qy	736	AAAGGAAGTGTCCCTCTGTGTGCAAGTGTGAAGTCTCCAGCTTTGCTCAGCTCTCCCA	795
Db	618	AAAGGAAGTGTCCCTCTGTGTGCAAGTGTGAAGTCTCAGCTTTGCTCAGCTCTCCCA	677
Qy	736	GACAGAAAAGCTGCGTCTTTTGTGTGTGGAACCCCAAGATCAAGAGATTTTGGAGCCC	855
Db	678	GACAGAAAAGCTGCGTCTTTTGTGTGTGGAACCCCAAGATCAAGAGATTTTGGAGCCC	737
Qy	856	GTGAAGAAAGAAAATGAGAGAGATGGGGACCTTGAACCTGAACGGGCACTTGTGTGTCGA	915
Db	738	GTGAAGAAAGAAAATGAGAGAGATGGGGACCTTGAACCTGAACGGGCACTTGTGTGTCGA	797
Qy	916	CAACCGCGTAGAATGATGCCAAACCGTCCACGAGGACGTCAAGACAGCGGCTGGGAAGCA	975
Db	798	CAACCGCGTAGAATGATGCCAAACCGTCCACGAGGACGTCAAGACAGCGGCTGGGAAGCA	857
Qy	976	GACAAAGATGAGAGAGACGCTGACATATCTGTGCAAGAACCTGTGACACGATCTCGGTG	103
Db	858	GACAAAGATGAGAGAGACGCTGACATATCTGTGCAAGAACCTGTGACACGATCTCGGTG	917
Qy	1036	AGTTTGACAGCCCTGTGATGACACGTTTGTGCGGGGCTGTGCTACGGGAGCTGGATGAGAGCG	109
Db	918	AGTTTGACAGCCCTGTGATGACACGTTTGTGCGGGGCTGTGCTACGGGAGCTGGATGAGAGCG	977
Qy	1096	TGTCCTCTGTGTCTTACCTGTGCGGCTGTGCCGTGAGAGCGGATCTGTAAAAACACATCTTC	115
Db	978	TGTCCTCTGTGTCTTACCTGTGCGGCTGTGCCGTGAGAGCGGATCTGTAAAAACACATCTTC	103
Qy	1156	AACAACTCTGTGGAAGCATCTTATCCAGGATCCAGACAAAGTGTGCGATGTAAGAAAT	121
Db	1038	AACAACTCTGTGGAAGCATCTTATCCAGGATCCAGACAAAGTGTGCGATGTAAGAAAT	109
Qy	1216	GTGCAAAAGTATGAGATGCAAGGAAATAAATCACTCAAGACATGCTGACGCCAAAGTCAG	127
Db	1098	GTGCAAAAGTATGAGATGCAAGGAAATAAATCACTCAAGACATGCTGACGCCAAAGTCAG	115
Qy	1276	CGGTCTTTTCTGATGAAGAAAGGAGTTCAAGAGGACTTGTGGAAGCTGTCAAGCGTTGAC	133
Db	1158	CGGTCTTTTCTGATGAAGAAAGGAGTTCAAGAGGACTTGTGGAAGCTGTCAAGCGTTGAC	121
Qy	1336	AGTGAAGTCTTCAAGATTAAGCAGGCAATACGTGTGTGTGCGGCAAGTGTCTTGAATACAA	139
Db	1218	AGTGAAGTCTTCAAGATTAAGCAGGCAATACGTGTGTGTGCGGCAAGTGTCTTGAATACAA	127
Qy	1396	AGGCAAGGAGGCGAGGCTCCCACTGCGCCAGGACCCGAGGGGAGGCAAGAGGCCCAAG	145
Db	1278	AGGCAAGGAGGCGAGGCTCCCACTGCGCCAGGACCCGAGGGGAGGCAAGAGGCCCAAG	133
Qy	1456	GCCCTGAGGAGATGCAACCTTCCACGTCCGTCAACCTGACGACAGCAGCAGTCCAGATTAACGTG	151
Db	1338	GCCCTGAGGAGATGCAACCTTCCACGTCCGTCAACCTGACGACGACAGCAGCAGTTCAGATTAACGTG	139
Qy	1516	TGCTCTTGTGCAAGAAAGCAGGCTCTGTGTGCACTGTGTCTTCAAGCCATGCCCCAGCG	157
Db	1398	TGCTCTTGTGCAAGAAAGCAGGCTCTGTGTGCACTGTGTCTTCAAGCCATGCCCCAGCG	145
Qy	1576	AGAGGAGAGCGGAGACAGGACCCGCGTGTGTGCGCCCTCAAGCAGTGTGCGGCTGTGCGCTGAG	163





```

Db      241 GGAAGATGATTAACAAGCTGAAGTTGTTAAGAACAGACATGCCCTTACAGACTGGG 300
Qy      391 GATGTCATCTACTTGGTGTACAGAAAGATGAACCGGAAACAACCTGGCATCTCTAT 450
Db      301 GATGTCATCTACTTGGTGTACAGAAAGATGAACCGGAAACAACCTGGCATCTCTAT 360
Qy      451 GAATCTTTAAGTGAAGAGAGAGCATGACAAAGATCCTTTGAAGCTAAACAAGAAAT 510
Db      361 GAATCTTTAAGTGAAGAGAGCATGACAAAGATCCTTTGAAGCTAAACAAGAAAT 420
Qy      511 GTGTTCCATGGGACCAAGATACCTCAGGTGCAAGTGCAGAGGCGAGGCGCATCCCGG 570
Db      421 GTGTTCCATGGGACCAAGATACCTCAGGTGCAAGTGCAGAGGCGAGGCGCATCCCGG 480
Qy      571 GTTCCTCCGTCCGCGCCCACTCAGGTGTGCTTTGAGAAACACAGCCATCAACATCG 630
Db      481 GTTCCTCCGTCCGCGCCCACTCAGGTGTGCTTTGAGAAACACAGCCATCAACATCG 540
Qy      631 AGTGAAGCTCTTCCCAAGCTGGGCTCTTCCAGAGAGCTTCTCTGACAGGCGA 690
Db      541 AGTGAAGCTCTTCCCAAGCTGGGCTCTTCCAGAGAGCTTCTCTGACAGGCGA 600
Qy      691 GAGCCTTCTCCAGATTGAGGATCTGAGGAGTGTGAGCATCTCCCTAAAGAAAGTGGTCCC 750
Db      601 GAGCCTTCTCCAGATTGAGGATCTGAGGAGTGTGAGCATCTCCCTAAAGAAAGTGGTCCC 660
Qy      751 TGTGTGGCAAGTGAAGTCTTCAGCTTTCCTCAGCTCTCCAGACAGAAAGACTGCG 810
Db      661 TGTGTGGCAAGTGAAGTCTTCAGCTTTCCTCAGCTCTCCAGACAGAAAGACTGCG 720
Qy      811 TCTCTTTTCTGCTGTGGAACCCAGAGTCAAGAGAAATTTGAAGCCCTGAAGAAAG 870
Db      721 TCTCTTTTCTGCTGTGGAACCCAGAGTCAAGAGAAATTTGAAGCCCTGAAGAAAG 780
Qy      871 AAGAGAGATGGGAGCTTGAACCTGAACGAGGCAAGTGTGTGTCGCAACAACGCGTAGAAAT 930
Db      781 AAGAGAGATGGGAGCTTGAACCTGAACGAGGCAAGTGTGTGTCGCAACAACGCGTAGAAAT 840
Qy      931 GCCCAACCGTCCACGAGAGACGTCAAGACAGCGGCTGGAGAGCCAGACAAAGTGAAGAG 990
Db      841 GCCCAACCGTCCACGAGAGACGTCAAGACAGCGGCTGGAGAGCCAGACAAAGTGAAGAG 900
Qy      991 AGCGTGAATGCATCATCTGCGCAGAGACTGCTGACAGACTGCTGAGTGTGACGCCCTGC 1050
Db      901 AGCGTGAATGCATCATCTGCGCAGAGACTGCTGACAGACTGCTGAGTGTGACGCCCTGC 960
Qy      1051 ATGCAACAGTTCTGGGCGGCTTGCTACTGGGCTGGATGGAAGCGTCGTCCGTGTGCT 1110
Db      961 ATGCAACAGTTCTGGGCGGCTTGCTACTGGGCTGGATGGAAGCGTCGTCCGTGTGCT 1020
Qy      1111 ACTCGCGCTGTCCCGTGAAGCGGATCTGTAAAAACAATCCTCAACAACCTCGTGA 1170
Db      1021 ACTCGCGCTGTCCCGTGAAGCGGATCTGTAAAAACAATCCTCAACAACCTCGTGA 1080
Qy      1171 GCATACCTCATCCAGCATCCAGCAAGAGTCCGAGTGAAGAGATGTGCAAGATATGAT 1230
Db      1081 GCATACCTCATCCAGCATCCAGCAAGAGTCCGAGTGAAGAGATGTGCAAGATATGAT 1140
Qy      1231 GCCAGGAATAAATCACTCAAGACATGCTGAGGCCCAAGTCAAGGAGGTCTTTTCTGAT 1290
Db      1141 GCCAGGAATAAATCACTCAAGACATGCTGAGGCCCAAGTCAAGGAGGTCTTTTCTGAT 1200
Qy      1291 GAAGAAGGAGTTCAGAGGACTGTGAGCTGTGACAGCTTGAAGAGTGTCTCAGAC 1350
Db      1201 GAAGAAGGAGTTCAGAGGACTGTGAGCTGTGACAGCTTGAAGAGTGTCTCAGAC 1260
Qy      1351 ATTAGCAAGCATAGTCTGTGTCGCGCAGTGTCTGAGTACAGAAAGCGAGCGCGCAG 1410
Db      1261 ATTAGCAAGCATAGTCTGTGTCGCGCAGTGTCTGAGTACAGAAAGCGAGCGCGCAG 1320
Qy      1411 CTTCCCACTGCGCCACACCCCGAGGCGAGCGCAGAGCCCAAGGCCCTGAGGAGATGCA 1470

```

```

Db      1321 CTTCCCACTGCGCCAGACCCGAGGCGAGCCAGAGGCCCAAGGCCCTGAGGAGTCA 1380
Qy      1471 CCTCCAGTCCGTGAGCTGACGACAGAGTCCAGATTAAGTGGCCCTGCAAGGA 1530
Db      1381 CCTCCAGTCCGTGAGCTGACGACAGAGTCCAGATTAAGTGGCCCTGCAAGGA 1440
Qy      1531 AGCCAGGCTGTGACCTGCTCTTCCAGGCCATATGCCGACCGGAGAGCGAGCGAG 1590
Db      1441 AGCCAGGCTGTGACCTGCTCTTCCAGGCCATATGCCGACCGGAGAGCGAGCGAG 1500
Qy      1591 CAGAACCCGCGTGTGCGCCCTCAGACATGTGCGGTCTGCTGACGCTTTTCTGCCACTG 1650
Db      1501 CAGAACCCGCGTGTGCGCCCTCAGACATGTGCGGTCTGCTGACGCTTTTCTGCCACTG 1560
Qy      1651 TACTGGGCTGACCCCGGACCGGCTGTACAGGCTGCTGCGCCGCTTTTGTGAGCTCAAC 1710
Db      1561 TACTGGGCTGACCCCGGACCGGCTGTACAGGCTGCTGCGCCGCTTTTGTGAGCTCAAC 1620
Qy      1711 CTGGGTGACAAAGTCTGGAAGGCTGTGAACAACAAGCTACAGATCAGACATCTCG 1770
Db      1621 CTGGGTGACAAAGTCTGGAAGGCTGTGTGAGAGCGCTGTGAACAACAAGCTACAGATCAGACATCTCG 1680
Qy      1771 AAGAAATTAACCTGGCAACCAAGGTTTGAATGAAACAAATGTTGACCGAGGCTCGTG 1830
Db      1681 AAGAAATTAACCTGGCAACCAAGGTTTGAATGAAACAAATGTTGACCGAGGCTCGTG 1740
Qy      1831 GCTTCCAGCGGAGAGTGTCTTGTCTGTGATTAACAAGTACCGGAGACACCGTTCTG 1890
Db      1741 GCTTCCAGCGGAGAGTGTCTTGTCTGTGATTAACAAGTACCGGAGACACCGTTCTG 1800
Qy      1891 TGTATGCTGTGAGGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGCAAACTT 1950
Db      1801 TGTATGCTGTGAGGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGCAAACTT 1860
Qy      1951 CTTGCTTCGAGTTCCAGTGGCCGTTAAATCCGCTCTGACTGCTAGTGGGCGGTAA 2010
Db      1861 CTTGCTTCGAGTTCCAGTGGCCGTTAAATCCGCTCTGACTGCTAGTGGGCGGTAA 1920
Qy      2011 TCCCGCACTCAGGTGAAGCTACACCGCATGAATTCATCATATCTGTAAACAGACA 2070
Db      1921 TCCCGCACTCAGGTGAAGCTACACCGCATGAATTCATCATATCTGTAAACAGACA 1980
Qy      2071 AGGTTCAAAAATAA 2085
Db      1981 AGGTTCAAAAATAA 1995

RESULT 6
ADM03388
ID ADM03388 Standard; cDNA; 2448 BP.
XX
AC ADM03388;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human cDNA of the invention SEQ ID NO:2073.
XX
KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
PN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
Yamamoto J, Isono Y, Hito Y, Otsuka K, Nagai K, Irie R, Tamechika I;

```

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI, 2003-723558/69.  
 DR P-PSDB; ADM05831.  
 XX  
 PT New polynucleotides and polypeptides are useful in gene therapy, for  
 PT developing a diagnostic marker or medicines for regulating their  
 PT expression and activity, or as a target of gene therapy.  
 XX  
 PS Claim 1, SEQ ID NO 2073; 305bp; English.  
 XX  
 CC The invention relates to a novel human polynucleotide and the encoded  
 CC polypeptide. A polynucleotide of the invention may have a use in gene  
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
 CC as a primer for synthesizing the polynucleotide or as a probe for  
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
 CC useful in gene therapy, for developing a diagnostic marker or medicines  
 CC for regulating their expression and activity, or as a target of gene  
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
 CC are useful as pharmaceutical agents. The present sequence represents a  
 CC cDNA sequence of the invention.  
 XX  
 SQ Sequence 2448 BP; 618 A; 624 C; 661 G; 545 T; 0 U; 0 Other;

Query Match 71.6%; Score 1919.4; DB 11; Length 2448;  
 Beft Local Similarity 88.5%; Pred. No. 0;  
 Matches 2227; Conservative 0; Mismatches 6; Indels 284; Gaps 2;

QY 218 GACGAGTGTGACCTTTCTTCCCGACGATTAACGTGCTCTGAGATCACTGTAGAA 277  
 DB 83 GACCGTGTGGACCTTCTCTCCCGAGATTAACGTGCTCTGAGATCACTGTAGAA 142  
 QY 278 TTGTGTGATGATTAATATGAGTCACTGAGTCACTGGAAGATCCAGACCGTGAACAG 337  
 DB 143 TTGTGTGATGATTAATATGAGTCACTGAGTCACTGGAAGATCCAGACCGTGAACAG 202  
 QY 338 TGATTAACAGCTGAAGTGTGTTAAGACGACATGCGCTTTACAGCTGGGGAATGCA 397  
 DB 203 TGATTAACAGCTGAAGTGTGTTAAGACGACATGCGCTTTACAGCTGGGGAATGCA 262  
 QY 398 TCTACTTGATGATGACGAGAAATGAACCGGAAACAACAGTGGCAATACCTCTATGATCTT 457  
 DB 263 TCTACTTGATGATGACGAGAAATGAACCGGAAACAACAGTGGCAATACCTCTATGATCTT 322  
 QY 458 TAAAGTAAAGCAAGGCAATGACACAAGATCTTTGAAAGCTTAAAGGAAAAATGTCTCC 517  
 DB 323 TAAAGTAAAGCAAGGCAATGACACAAGATCTT----- 356  
 QY 518 ATGGGACCAAGATATCTCAGGTGACAGGTGACGGGCGAGCGGCGATCCCGGGTCCCTC 577  
 DB 357 ----- 356  
 QY 578 CGTCGTCCCGGCACTCAGGTGTGCTTTGAGGAAACAAGCANTCAATGACGTGAG 637  
 DB 357 ----- 356  
 QY 638 AACTCTTCCCAAGACCTCGGCTCTTCCAGAGACCTTCTCTGAGGGGAGAGGCTT 697  
 DB 357 ----- 356  
 QY 698 CCTCAGATTGGGCTGCGGGGTGGTGGCATCTCCCTAAAGGAAGGTGCTCTGAG 757  
 DB 357 -----TGGGCTCTGGGGGTGGTGGCATCTCCCTAAAGGAAGGTGCTCTGAG 406  
 QY 758 CAAGATGAAGTCTCAGCTTTGCTCAGCTCTTCCAGACAGAAAGAATCGCTCTTTT 817  
 DB 407 CAAGATGAAGTCTCAGCTTTGCTCAGCTCTTCCAGACAGAAAGAATCGCTCTTTT 466  
 QY 818 CGTGTGTTGAACCCAGGATCAGAGGATTTGGAGCCCGTGAAGAAAAATGAGAGAG 877  
 DB 467 CGTGTGTTGAACCCAGGATCAGAGGATTTGGAGCCCGTGAAGAAAAATGAGAGAG 526  
 QY 878 ATGGGACCTTGACCTGAACGGGCGAGTTGTTGGTGCACAAACGGGTAGAAATGCCCAA 937

DB 527 ATGGGACCTTGACCTGAACGGGCGAGTTGTTGGTGCACAAACCGGTAGAAATGCCCAA 586  
 QY 938 CCGTCCACGAGAGCTGACAGACAGCGGCTGGGAAACCAACAAATGAGAGAGAGCTGA 997  
 DB 587 CCGTCCACGAGAGAGTGAAGACAGCGGCTGGGAAACCAACAAATGAGAGAGAGCTGA 646  
 QY 998 CATGATCATATGTCGAGAGCTGCTGACAGCTGCGTGAAGTTTGAAGCCCTGATGACA 1057  
 DB 647 CATGATCATATGTCGAGAGCTGCTGACAGCTGCGTGAAGTTTGAAGCCCTGATGACA 706  
 QY 1058 CGTTCTGCGGGCTTGGTAACTCGGGCTGAGATGAGCGCTGCTGCTGCTGCTGCTGCTG 1117  
 DB 707 CGTTCTGCGGGCTTGGTAACTCGGGCTGAGATGAGCGCTGCTGCTGCTGCTGCTGCTG 766  
 QY 1118 GCTGTCCGTTGAGACCGGATCTGTAAAAACAATCTCTCAACAACTGCTGAGACATACC 1177  
 DB 767 GCTGTCCGTTGAGACCGGATCTGTAAAAACAATCTCTCAACAACTGCTGAGACATACC 826  
 QY 1178 TCATCCAGCATCCAGACAGAGTGCAGTGAAGAAATGTGCAAAATGTGATGCCAGGA 1237  
 DB 827 TCATCCAGCATCCAGACAGAGTGCAGTGAAGAAATGTGCAAAATGTGATGCCAGGA 886  
 QY 1238 ATAAATCACTCAAGACATGCTGAGGCCCAAAGTCAGGCGCTTTTCTGATGAAGAG 1297  
 DB 887 ATAAATCACTCAAGACATGCTGAGGCCCAAAGTCAGGCGCTTTTCTGATGAAGAG 946  
 QY 1298 GAGATTCAGAGACCTGCTGAGCTGTGACAGTTTGAACATGATGCTCTGACAGATTAAG 1357  
 DB 947 GAGATTCAGAGACCTGCTGAGCTGTGACAGTTTGAACATGATGCTCTGACAGATTAAG 1006  
 QY 1358 AGCCATACGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1417  
 DB 1007 AGCCATACGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066  
 QY 1418 ACTGCCAGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1477  
 DB 1067 ACTGCCAGACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1126  
 QY 1478 CGTCCGTGAGCTGACAGACAGAGTCCAGAGATTAAGTGTGCTCTGCTGCAAGAAAGCA 1537  
 DB 1127 CGTCCGTGAGCTGACAGACAGAGTCCAGAGATTAAGTGTGCTCTGCTGCAAGAAAGCA 1186  
 QY 1538 CCTGTGCACTGTGCTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1597  
 DB 1187 CCTGTGCACTGTGCTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1246  
 QY 1598 CGCGTGTGCGCCCTCAGACAGTGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1657  
 DB 1247 CGCGTGTGCGCCCTCAGACAGTGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306  
 QY 1658 GCTGCAACCGGACCGGCTGTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1701  
 DB 1307 GCTGCAACCGGACCGGCTGTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1366  
 QY 1702 -----GAGCTCAA 1709  
 DB 1367 GCTTTCCTGCGACATCAATTAATAACAGTAAATGCTCCCTTCAATTCGCGAGAGCTCAA 1426  
 QY 1710 CTTGGGTGACAAAGTGTCTGAGACCGGCTGTAAGCAACAAAGTGTCTGAGTCAATCTCT 1769  
 DB 1427 CTTGGGTGACAAAGTGTCTGAGACCGGCTGTAAGCAACAAAGTGTCTGAGTCAATCTCT 1486  
 QY 1770 GAAGAAATTAATCTGCAACAGAGGTTTGAATGAGAAAAATGATGTAAGAGAGCTGCT 1829  
 DB 1487 GAAGAAATTAATCTGCAACAGAGGTTTGAATGAGAAAAATGATGTAAGAGAGCTGCT 1546  
 QY 1830 GGTCTTCCACGCGGAGATGTTTCTGCTGTCTGATTAAGAGTCAACGAGGAAACACCTGCT 1889  
 DB 1547 GGTCTTCCACGCGGAGATGTTTCTGCTGTCTGATTAAGAGTCAACGAGGAAACACCTGCT 1606  
 QY 1890 GTGTTAATGCTGTGAGCTGCGAGCTTCCGTGAGCTGACCTATCAGATGCGGAGAAAT 1949



Db 1425 TACTGCTGCGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATGGAGCAATTCCT 1366  
Qy 1954 GCTTCGAGATTGCGAGTGGCCGTATACATCCGCTGACTGCTAGTGGGCGGTAACTGC 2013  
Db 1365 GCTTCGAGATTGCGAGTGGCCGTATACATCCGCTGACTGCTAGTGGGCGGTAACTGC 1306  
Qy 2014 CGCAGCTCAGGTGAGTGAAGCTCAGCAGCCATGAAATTCATATCTGTGAACAGACAG 2073  
Db 1305 CGCAGCTCAGGTGAGTGAAGCTCAGCAGCCATGAAATTCATATCTGTGAACAGACAG 1246  
Qy 2074 TTCAAAAATGAGATCCAGAGGCGCTGAGAGCTTTCAGCAGTGAAGTGAAGAGAGG 2133  
Db 1245 TTCAAAAATGAGATCCAGAGGCGCTGAGAGCTTTCAGCAGTGAAGTGAAGAGAGG 1186  
Qy 2134 TGTTTTAAATACAGAGACAGACAGTGAAGTGTTCACAGCCCTGAGGAGAGG 2193  
Db 1185 TGTTTTAAATACAGAGACAGACAGTGAAGTGTTCACAGCCCTGAGGAGAGG 1126  
Qy 2194 AGCAGAGGCTCTCGACAGAGTCTGCGGCTGACTCTTCTGTGAGAGCTTTTACCTCTGA 2253  
Db 1125 AGCAGAGGCTCTCGACAGAGTCTGCGGCTGACTCTTCTGTGAGAG--TTTACCTCTGA 1068  
Qy 2254 GTGAGACCTCCCGAGAGCCCGGGGGCGGAGCCCGCTGCTGAGAGGCTGGGAG 2313  
Db 1067 GTGAGACCTCCCGAGAGCCCGGGGGCGGAGCCCGCTGCTGAGAGGCTGGGAG 1008  
Qy 2314 GGCCTGTGTGATGAGATCAGAGACAGACAGACAGTCTTCTGTAACATGCGGCGCTCCGC 2373  
Db 1007 GGCCTGTGTGATGAGATCAGAGACAGACAGACAGTCTTCTGTAACATGCGGCGCTCCGC 948  
Qy 2374 GAGAGGGGAGTGTGCTCTTTTGTACATTTTCCGAAATCAGTAAAGCGAAGTCTG 2433  
Db 947 GAGAGGGGAGTGTGCTCTTTTGTACATTTTCCGAAATCAGTAAAGCGAAGTCTG 888  
Qy 2434 TTTTCAGAAAAGTTCAAGGAGAGAGGAGCAATTTCAAAAACATTTTCAAGAGA 2493  
Db 887 TTTTCAGAAAAGTTCAAGGAGAGAGGAGCAATTTCAAAAACATTTTCAAGAGA 828  
Qy 2494 GGGAGCATTAAGTTTACAGCTTACAGAGCTGACATATCTGCTGCGGAGAACACA 2553  
Db 827 GGGAGCATTAAGTTTACAGCTTACAGAGCTGACATATCTGCTGCGGAGAACACA 768  
Qy 2554 GGAATTTATCTATTTTATTTTATAGTGTGCTTATCTTCTATTAAGATTAAAT 2613  
Db 767 GGAATTTATCTATTTTATTTTATAGTGTGCTTATCTTCTATTAAGATTAAAT 708  
Qy 2614 GTCAAAAATGAGCAAAATATATATTTTATTAATTTACAAATGACAAA 2666  
Db 707 GTCAAAAATGAGCAAAATATATATTTTATTAATTTACAAATGACAAA 655  
RESULT 8  
AAS25843  
ID AAS25843 strand; cDNA; 1311 BP.  
XX  
AC AAS25843;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 22.  
XX  
XX Human; immunosuppressive; antiarthritic; ssr; antirheumatic; cyostatic;  
cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective;  
antibacterial; virucide; fungicide; ophthalmological; vulnery;  
secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
cerebral ischemia; angiogenesis; nervous system disorder;  
Alzheimer's disease; infection; ocular disorder; corneal infection;  
wound healing; epithelial cell proliferation; skin ageing; food additive;  
preservative; antiproliferative.  
XX  
OS Homo sapiens.

XX  
PN W020015322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001341.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186356P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209457P.  
PR 28-JUN-2000; 2000US-0214866P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.



PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241211P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-024617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250310P.  
PR 01-DEC-2000; 2000US-0250311P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251866P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX  
PA  
XX

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SW;  
XX  
XX MPI; 2001-489783/53.  
DR P-PSDB; AUI15856.  
PT New nucleic acid molecules encoding 461 human secreted proteins for  
PR diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.  
XX  
XX Claim 1; SEQ ID NO 22; 980pp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed

Query Match 35.3%; Score 945; DB 4; Length 1311;  
Best Local Similarity 81.1%; Pred. No. 6, 1e-217; Indels 282; Gaps 4;  
Matches 1267; Conservative 6; Mismatches 7; Indels 282; Gaps 4;

QY 56 TCCGGGTTTCGGCGCGGGGAGATGTGATCCGATGAGCGGCCGAGAAAGCAAC 115  
Db 5 TCCGGGTTTCGGCGCGGGGAGATGTGATCCGATGAGCGGCCGAGAAAGCAAC 64  
QY 116 AGTCGCGCGCGCGGACCCCTTGGGAGCGCTCTTGGCGCGCGGAGGCGGAGC 175  
Db 65 AGTCGCGCGCGCGGAGCCCTTGGGAGCGCTCTTGGCGCGCGGAGGCGGAGC 124  
QY 176 CGCAGCTCTCTGAGGAAGCGGAGTGGACCATGGGCGGAGAGGTGCACCTT 235  
Db 125 CGCAGCTCTCTGAGGAAGCGGAGTGGACCATGGGCGGAGAGGTGCACCTT 184  
QY 236 CCTTCCCGCAGATTAATCTGTCTGAGATCACTGTAGATTTAGTGAATAAT 295  
Db 185 CCTTCCCGCAGATTAATCTGTCTGAGATCACTGTAGATTTAGTGAATAAT 244  
QY 296 CAGGTGAGTGAACCTGGAATACAGACCAAGTGAACAGTGAATTAACCTGAAG 355  
Db 245 CAGGTGAGTGAACCTGGAATACAGACCAAGTGAACAGTGAATTAACCTGAAG 304  
QY 356 TTGTTAAGAGAGACATGCCCTTTACAGATCGGGGATGCTATCTTGGTGCAGGA 415  
Db 305 TTGTTAAGAGAGACATGCCCTTTACAGATCGGGGATGCTATCTTGGTGCAGGA 364  
QY 416 AGAATGAACCGGAACCAACGTGAGTACCTTATGAATCTTTAAGTGAAGGAAGCA 475  
Db 365 AGAATGAACCGGAACCA----- 382  
QY 476 TGAACAGAAATCTTTGAAGCTTAACAGAAATGTGTTCCATGGACCAAGATCCT 535  
Db 383 ----- 382  
QY 536 CAGGTGAGTGAAGGCGGAGGCGGATCCCGGGTCCCTCGTGTGCGCGGCACTC 595

[illegible]

Db	1288	CA	1285
RESULT 9			
ABX73184			
ID	ABX73184	standard; DNA; 1311 BP.	
XX			
AC	ABX73184;		
XX			
DT	18-MAR-2003	(first entry)	
DE			
XX			
Human novel polynucleotide #12.			
XX			
OS	Homo sapiens.		
XX			
XX	US2002132753-A1.		
PD	19-SEP-2002.		
XX			
PF	17-JAN-2001; 2001US-00764864.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216809P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	22-AUG-2000; 2000US-0226668P.		
PR	30-AUG-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0229509P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	08-SEP-2000; 2000US-0231413P.		
PR	21-SEP-2000; 2000US-0234223P.		
PR	21-SEP-2000; 2000US-0234274P.		
PR	25-SEP-2000; 2000US-0234977P.		
PR	27-SEP-2000; 2000US-0235834P.		
PR	29-SEP-2000; 2000US-0235327P.		
PR	29-SEP-2000; 2000US-0236367P.		
PR	29-SEP-2000; 2000US-0236368P.		
PR	29-SEP-2000; 2000US-0236369P.		
PR	29-SEP-2000; 2000US-0236370P.		
PR	02-OCT-2000; 2000US-0236802P.		
PR	02-OCT-2000; 2000US-0237037P.		
PR	02-OCT-2000; 2000US-0237038P.		
PR	02-OCT-2000; 2000US-0237039P.		
PR	02-OCT-2000; 2000US-0237040P.		
PR	13-OCT-2000; 2000US-0239935P.		
PR	20-OCT-2000; 2000US-0240960P.		
PR	20-OCT-2000; 2000US-0241785P.		

PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 DR MPI; 2003-147444/14.  
 DR P-PSDB; ABUS4924.  
 XX  
 PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
 PT inhibiting or preventing e.g. neural, immune system, muscular,  
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
 PT renal disorders.  
 XX  
 PS Claim 1; SEQ ID NO 22; 402pp; English.  
 XX  
 CC The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
 CC leukemia), hyperproliferative disorders (e.g. Hodgkin's disease and  
 CC leukemias), inflammatory diseases (e.g. septic shock, dermatitis and  
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
 CC human novel polynucleotides of the invention  
 CC  
 SO Sequence 1311 BP; 306 A; 359 C; 399 G; 239 T; 0 U; 8 Other;  
 Query Match 35.3%; Score 945; DB 8; Length 1311;  
 Best Local Similarity 81.1%; Pred. No. 6.1e-217;  
 Matches 1267; Conservative 6; Mismatches 7; Indels 282; Gaps 4;

QY 476 TGCACAGAAATCTTTGAAGCTAACAGAAAAATGTTCATGAGCAACAAATACCT 535  
 DB 383 ----- 382  
 QY 536 CAGGTGACAGTGCAGGGCGAGGGCCGATCCCCGGGTCCCTCGTCGCCGCCACATC 595  
 DB 383 ----- 382  
 QY 596 AGGTGCTTTGAGAAACCAAGCCATCAACATCGACGTCAGACCTCTTCCCAAGCCT 655  
 DB 383 ----- 382  
 QY 656 CGGCTCTTCCACGAGCCTTCTCTGACGGCGAGAGGTTCTTCAGTTGAGGTCTG 715  
 DB 383 -----GGTCTG 388  
 QY 716 GGGGTGTGGCATCTCCCTTAAGAGAGTGTCTCTGTGTGCAAGTGAAGTCTCA 775  
 DB 389 GGGGTGTGGCATCTCCCTTAAGAGAGTGTCTCTGTGTGCAAGTGAAGTCTCA 448  
 QY 776 GCTTTGCTCAGCTCTCCAGACAGAAAGACTGCGTCTTTCTGTCTGTGAACCCCAAG 835  
 DB 449 GCTTTGCTCAGCTCTCCAGACAGAAAGACTGCGTCTTTCTGTCTGTGAACCCCAAG 508  
 QY 836 ATCAGAGGATTTGAGACCCGTGAAGAAATGAGAGAGATGGAGACTTGAACCTGA 895  
 DB 509 ATCAGAGGATTTGAGACCCGTGAAGAAATGAGAGAGATGGAGACTTGAACCTGA 568  
 QY 896 ACGGGAGTTGTTGTGTGCAACACCGCGTAGAAATGCCAAACCGTCCACAGAGAGTCA 955  
 DB 569 ACGGGAGTTGTTGTGTGCAACACCGCGTAGAAATGCCAAACCGTCCACAGAGAGTCA 628  
 QY 956 GAGCAGCGGCTGGGAAGCCAGCAAGATGAGAGAGACGTGACATGATCATCTGCAAG 1015  
 DB 629 GAGCAGCGGCTGGGAAGCCAGCAAGATGAGAGAGACGTGACATGATCATCTGCAAG 688  
 QY 1016 ACCTGTGACACACTGCGTAGATTGTCAGACCCCTGACATGACAGTCTTCTCGGCGTGTCT 1075  
 DB 689 ACCTGTGACACACTGCGTAGATTGTCAGACCCCTGACATGACAGTCTTCTCGGCGTGTCT 748  
 QY 1076 ACTCGGCTGATGATGAGACCGCTGTCTGTCTTCTTCTGATGAAGAGGATTCAGAGACCTTC 1135  
 DB 749 ACTCGGCTGATGATGAGACCGCTGTCTGTCTTCTTCTGATGAAGAGGATTCAGAGACCTTC 808  
 QY 1136 TCTGTAAAAACCACTCTCTCAACAACCTGTGTGAAGCATCTCATCCAGATCCAGACA 1195  
 DB 809 TCTGTAAAAACCACTCTCTCAACAACCTGTGTGAAGCATCTCATCCAGATCCAGACA 868  
 QY 1196 AGAGTGCAGTGAAGAAAGATGTGCAAAAGTATGATGCGCAGAAATCACTCAAGACA 1255  
 DB 869 AGAGTGCAGTGAAGAAAGATGTGCAAAAGTATGATGCGCAGAAATCACTCAAGACA 928  
 QY 1256 TGTCTCAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGATTCAGAGACCTTC 1315  
 DB 929 TGTCTCAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAGGATTCAGAGACCTTC 988  
 QY 1316 TGAAGCTGTCAAGCTTTGACATGATGCTCTCAACATTAAGCAGCATACGTGTGTGCC 1375  
 DB 989 TGAAGCTGTCAAGCTTTGACATGATGCTCTCAACATTAAGCAGCATACGTGTGTGCC 1048  
 QY 1376 GGCAGTGTCTGATGTACAGAAAGGAGGGGCGCAGCTCCCACTGCGCCAGACCCGAGG 1435  
 DB 1049 GGCAGTGTCTGATGTACAGAAAGGAGGGGCGCAGCTCCCACTGCGCCAGACCCGAGG 1108  
 QY 1436 GCGAGCCAGAGACCCCAAGAGCCTTGGGGATGCAACCTTCAAGTCCGTCAAGCTTGAACA 1495  
 DB 1109 GCGAGCCAGAGACCCCAAGAGCCTTGGGGATGCAACCTTCAAGTCCGTCAAGCTTGAACA 1168  
 QY 1496 CAGCAGTCAAGATTAAGTGTCCCTCTGTGCAAGAGGACAGCCCTGTGTCAACTGTCTGT 1555  
 DB 1169 CAGT-----CAGGATTAAGTGTCCCTCTGTGCAAGAGGACAGCCCTGTGTCAACTGTCTGT 1224  
 QY 1556 TCCAGCCATGCCCCAGACCGAGAGGCGGAGACAGACA-----CCGCGTGTGCCCTCTGAG 1614



PR 09-JUN-2000; 2000JP-00241899.  
 XX (HELI-) HELIX RES INST.  
 XX  
 PA Oca T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX  
 PR Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX  
 PS Claim 1; SEQ ID NO 3663; 2537pp + Sequence Listing; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 CC  
 XX Sequence 816 BP; 182 A; 216 C; 253 G; 162 T; 0 U; 3 Other;  
 XX  
 SQ  
 Query Match 27.1%; Score 727; DB 4; Length 816;  
 Best Local Similarity 93.5%; Pred. No. 1.5e-164;  
 Matches 793; Conservative 0; Mismatches 18; Indels 37; Gaps 2;

QY 437 TGGCATACCTCTATGATATCTTTAAGTAAACCAAGCATGACACAGATCTTTGAAG 496  
 DB 425 TGGCATACCTCTATGATATCTTTAAGTAAACCAAGCATGACACAGATCTTT 480  
 QY 497 CTAAACAGAAAATGTGTTCCATGGGACCAAGATACCTCAGGTGAGTGAAGGCGAG 556  
 DB 481 -----GATACCTCAGGTGAGTGAAGGCGAG 508  
 QY 557 GGGCCGATCCCGGGTCCCTCCGTGTGCGCCGACATCAGGTGCTTTAGGAACAC 616  
 DB 509 GGGCCGATCCCGGGTCCCTCCGTGTGCGCCGACATCAGGTGCTTTAGGAACAC 568  
 QY 617 ACCCATCAACATGACATGACATCTTCCGACAGCTCGGCTCTTCCAGGAGCTT 676  
 DB 569 ACCCATCAACATGACATGACATCTTCCGACAGCTCGGCTCTTCCAGGAGCTT 628  
 QY 677 CTCTGCAAGGAGGAGAGCGTTCTTCAGTTGGGTCTGGGGGTGGGATCTCCCTTA 736  
 DB 629 CTCTGCAAGGAGGAGAGCGTTCTTCAGTTGGGTCTGGGGGTGGGATCTCCCTTA 688  
 QY 737 AAGGAAGTGTCTCTGTGGCAAGTGAAGTCTCCAGCTTTGCTTCCAGCTTCCAG 796  
 DB 689 AAGGAAGTGTCTCTGTGGCAAGTGAAGTCTTCAGCTTTGCTTCCAGCTTCCAG 748  
 QY 797 ACAGAAAGCTGCTCTTCTGTGTTGAACCCAGATCAGAGGA-TTTGAGAGCC 855  
 DB 749 ACAGAAAGCTGCTCTTCTGTGTTGAACCCAGATCAGAGGA-TTTGAGAGCC 808  
 QY 856 GTGAAGAA 863  
 DB 809 GTGAAGAA 816

RESULT 12  
 AAH06846  
 ID AAH06846 standard; cDNA; 824 BP.  
 XX  
 AC AAH06846;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DB Human cDNA clone (5'-primer) SEQ ID NO:3681.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-00116126.  
 XX  
 PR 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Oca T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX  
 PR Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 PS Claim 1; SEQ ID NO 3681; 2537pp + Sequence Listing; English.  
 XX





PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234233P.  
PR 21-SEP-2000; 2000US-0234234P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234598P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241855P.  
PR 20-OCT-2000; 2000US-0241865P.  
PR 20-OCT-2000; 2000US-0241877P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246809P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0253678P.  
  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-489783/53.  
DR F-PSDB; AAIJ6317.  
XX  
PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives.  
XX  
XX  
PS Claim 1, SEQ ID NO 463; 980pp; English.  
XX  
XX  
CC The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Antibodies to the proteins can also be used in  
CC alleviating symptoms associated with the disorders and in diagnostic  
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
CC diseases e.g. Rheumatoid arthritis, hyperproliferative disorders e.g.  
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac  
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,  
CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
CC and many other disorders listed in the specification. The polypeptides  
CC can also be used to aid wound healing and epithelial cell proliferation,  
CC to prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage

CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention. Note: The  
CC sequence data for this patent did not form part of the printed

Query Match 20.0%; Score 536.6; DB 4; Length 693;  
Best Local Similarity 92.5%; Pred. No. 9,8e-119;

Matches 608; Conservative 7; Mismatches 13; Indels 29; Gaps 4;

```
QY 1044 GCCCTGATGACACGCTTCTGCGCGGCTTGTACTGCGGCTGATGAGAGCGTCCCT 1103
DB 14 GCCCTGATGACACGCTTCTGCGCGGCTTGTACTGCGGCTGATGAGAGCGTCCCT 73
QY 1104 GGTGCTACTGCGCGGCTGCGGCTGAGAGGAGTCTGTAATAACCATCTCTCAACACT 1163
DB 74 GGTGCTACTGCGCGGCTGCGGCTGAGAGGAGTCTGTAATAACCATCTCTCAACACT 133
QY 1164 CGTGAAGCATACCTCATCCAGCATCCAGCAAGAGTCCAGTGAAGAAGATGTCAAG 1223
DB 134 CGTGAAGCATACCTCATCCAGCATCCAGCAAGAGTCCAGTGAAGAAGATGTCAAG 193
QY 1224 TATGATGCGAGAAATAAATCACTCAGACATGCTGACGCCAAAGTCAAGCGGCTCTT 1283
DB 194 TATGATGCGAGAAATAAATCACTCAGACATGCTGACGCCAAAGTCAAGCGGCTCTT 253
QY 1284 TTTGATGAGAAAGGAGTTGAGAGGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTG 1343
DB 254 TTTGATGAGAAAGGAGTTGAGAGGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTG 313
QY 1344 CTGACAGATTAGCCAGCCATACGTGCTGTGCGGAGAGTCTGCTGAGTACAGAGCAGAGC 1403
DB 314 CTGACAGATTAGCCAGCCATACGTGCTGTGCGGAGAGTCTGCTGAGTACAGAGCAGAGC 373
QY 1404 GCGCGAGCTCTCCCACTGCGCCAGCACCGAGGCGAGAGCCAGAGCCCAAGGCTTGG 1463
DB 374 GCGCGAGCTCTCCCACTGCGCCAGCACCGAGGCGAGAGCCCAAGGCTTGG 433
QY 1464 GGATGACCCCTCCAGCTGCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTG 1523
DB 434 GGATGACCCCTCCAGCTGCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTG 489
QY 1524 GCAAGAGAGCCAGCGCTGTGTGCACTGTCTTCCAGCCATGCCGAGCCGAGAGCGGA 1583
DB 490 GCAAGAGAGCCAGCGCTGTGTGCACTGTCTTCCAGCCATGCCGAGCCGAGAGCGGA 548
QY 1584 GGGCAGCAGAGCCCGCGTGTGCGCCCTCAGCAGTGTGGGCTGTGCTGACGCTTTCTG 1643
DB 549 -CGCAGCAGAGCCCGCGTGTGCGCCCTCAGCAGTGTGGGCTGTGCTGACGCTTTCTG 584
QY 1644 CCACCTGATCTGAGGCTGTGACCCGAGACCGGCTGTACAGGCTGCTGAGCCGCTTTTG 1700
DB 585 CCACCTGATCTGAGGCTGTGACCCGAGACCGGCTGTACAGGCTGCTGAGCCGCTTTTG 641
```

RESULT 14

ID ABX73645  
ABX73645 standard; DNA; 693 BP.

AC ABX73645;

DT 18-MAR-2003 (first entry)

DE Human novel polynucleotide #473.

XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;

XX muscular disorder; respiratory disease; reproductive disorder;

XX gastrintestinal disorder; pulmonary disorder; cardiovascular disorder;

XX hyperproliferative disorder; inflammatory disease; allergic reaction;

XX blood related disorder; cancer; immunosuppressive; antiinflammatory;

XX cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;

XX haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX US2002132753-A1.

XX 19-SEP-2002.

XX 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216647P.

XX 11-JUL-2000; 2000US-0217487P.

XX 14-JUL-2000; 2000US-0217965P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 22-AUG-2000; 2000US-0226868P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 05-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 08-SEP-2000; 2000US-0229513P.

XX 21-SEP-2000; 2000US-0231413P.

XX 25-SEP-2000; 2000US-0234274P.

XX 27-SEP-2000; 2000US-0234997P.

XX 29-SEP-2000; 2000US-0236327P.

XX 29-SEP-2000; 2000US-0236367P.

XX 29-SEP-2000; 2000US-0236368P.

XX 29-SEP-2000; 2000US-0236369P.

XX 02-OCT-2000; 2000US-0236802P.

XX 02-OCT-2000; 2000US-0237037P.

XX 02-OCT-2000; 2000US-0237038P.

XX 02-OCT-2000; 2000US-0237039P.

XX 13-OCT-2000; 2000US-0239350P.

XX 20-OCT-2000; 2000US-0240960P.

XX 20-OCT-2000; 2000US-0241785P.

XX 01-NOV-2000; 2000US-0241809P.

XX 17-NOV-2000; 2000US-0244617P.

XX 08-DEC-2000; 2000US-0251856P.

XX 08-DEC-2000; 2000US-0251868P.

XX 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

XX (RUBEN/) RUBEN S M.

XX (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2003-147444/14.

XX P-PSDB; ABUS5385.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,

XX inhibiting or preventing e.g. neural, immune system, muscular,

XX respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or

XX renal disorders.

XX Claim 1; SEQ ID NO 483; 402pp; English.

XX The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
 CC nasal vasculitis), nasal polyps and sinusitis), reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
 CC leukemia), inflammatory reactions and conditions (e.g. asthma, blood  
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
 CC human novel polynucleotides of the invention

Sequence 693 BP; 145 A; 214 C; 197 G; 123 T; 0 U; 14 Other;

Query Match 20.0%; Score 536.6; DB 8; Length 693;  
 Best Local Similarity 92.5%; Pred. No. 9.8e-119;  
 Matches 608; Conservative 7; Mismatches 13; Indels 29; Gaps 4;

1044 GCCCTGATGACACACGTTCTGCGGCTTGTACTGCGGCTGATGAGAGCGCTCTCCCT 1103  
 14 GCCCTGATGACACACGTTCTGCGGCTTGTACTGCGGCTGATGAGAGCGCTCTCCCT 73  
 1104 GTGTCTTACCTGCGCTGTCCTGCGGATCTGTAAACCAACATCTCTCAACACT 1163  
 74 GTGTCTTACCTGCGCTGTCCTGCGGATCTGTAAACCAACATCTCTCAACACT 133  
 1164 CGTGAAGATACCTATCCAGATCCAGACAGAGTCCGAGTGAAGAGTGAAG 1223  
 134 CGTGAAGATACCTATCCAGATCCAGACAGAGTCCGAGTGAAGAGTGAAG 193  
 1224 TATGATGCGAGGATTAATCACTCAAGATCTGTCAGCCCAAGTGGCGCTCTT 1283  
 194 TATGATGCGAGGATTAATCACTCAAGATCTGTCAGCCCAAGTGGCGCTCTT 253  
 1284 TTCTGATGAAGAGGATTCAGAGAGCTGTGAGAGCTGTCAAGCTTGAAGTGC 1343  
 254 TTCTGATGAAGAGGATTCAGAGAGCTGTGAGAGCTGTCAAGCTTGAAGTGC 313  
 1344 CTCAGACATTAAGCAGCCATAGCTGTCGTCGCGGAGTGTCTGATACAGAGGAGC 1403  
 314 CTCAGACATTAAGCAGCCATAGCTGTCGTCGCGGAGTGTCTGATACAGAGGAGC 373  
 1404 GCGGACAGCTCCCACTGCGCAGACACCGGAGGCGAGAGCCCAAGGCTCTGG 1463  
 374 GCGGACAGCTCCCACTGCGCAGACACCGGAGGCGAGAGCCCAAGGCTCTGG 433  
 1464 GGATGACCTCTCCAGCTCGTCAGCTTGAAGAGAGATTCAGTGTGCTCTCT 1523  
 434 GGATGACCTCTCCAGCTCGTCAGCTTGAAGAGAGATTCAGTGTGCTCTCT 489  
 1524 GGAAAGAGCAGCCCTGTCGACCTGTCCTTCAAGCCATGTCGAGAGAGAGC 1583  
 490 GGAAAGAGCAGCCCTGTCGACCTGTCCTTCAAGCCATGTCGAGAGAGAGC 548  
 1584 GCGGACAGCAGCCCGCTGTCGCTTGAAGAGTGTGCTGCTGAGAGCTTCTG 1643  
 549 GCGGACAGCAGCCCGCTGTCGCTTGAAGAGTGTGCTGCTGAGAGCTTCTG 584  
 1644 CCACTGTACTGAGGCTGACCCGAGACCGGCTGCTACGAGTGTGCTGAGCCCGTTG 1700  
 585 CCACTGTACTGAGGCTGACCCGAGACCGGCTGCTACGAGTGTGCTGAGCCCGTTG 641

XX 21-AUG-2000 (first entry)  
 XX Human secreted expressed sequence tag SEQ ID NO:911.  
 DE Human secreted expressed sequence tag SEQ ID NO:911.  
 XX Human; mouse; chicken; rat; secreted expressed sequence tag; EST;  
 XX expressed sequence tag; EST; probe; chemotactic; proliferative;  
 XX immunomodulatory; haematopoietic; chemokinetic; angiogenic; haemostatic;  
 XX thrombolytic; antiinflammatory; cyrostatic; antibacterial; antifungal;  
 XX antiviral; antidiabetic; antiaethmatic; vulnary; antiparkinsonian;  
 XX antitumor; osteopathic; neuroprotective; nootropic; antiporiatic;  
 XX cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;  
 XX autoimmune disorder; multiple sclerosis; allergic condition;  
 XX insulin dependent diabetes; aschma; myeloid cell deficiency; ulcer;  
 XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 XX central nervous system disorder; Alzheimer's disease; stroke;  
 XX Parkinson's disease; Huntington's disease; coagulation disorder;  
 XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
 XX infection; depression; psoriasis; ss.

OS Homo sapiens.  
 XX  
 XX MO200021991-A1.  
 XX  
 XX 20-APR-2000.  
 XX  
 XX 15-OCT-1999; 99MO-US024206.  
 XX  
 XX 15-OCT-1998; 98US-0104436P.  
 XX  
 XX (GENW) GENETICS INST INC.  
 XX  
 XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 XX Merberg D, Treacy M, Bowman MR;  
 XX WPI; 2000-317938/27.  
 XX  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (ESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders.  
 XX  
 XX Claim 1; Page 437; 803pp; English.

AA44326 to AA445925 represent specifically claimed secreted expressed  
 CC sequence tags (ESTs), isolated from human, mouse, chicken and rat tissue  
 CC sources. The ESTs can have a range of activities depending on the  
 CC tissues they were isolated from. The activities include: Chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; antiinflammatory; cyrostatic; antibacterial;  
 CC antifungal; antiviral; antidiabetic; antiaethmatic; vulnary; antitumor;  
 CC osteopathic; neuroprotective; nootropic; antiparkinsonian; antiporiatic;  
 CC cerebroprotective; anticonvulsant; and antidepressant. The ESTs can be  
 CC used for gene therapy and in vaccines. The ESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the ESTs. Proteins encoded by the ESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis), insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AA445926 to AA445931 represent linker variants which are given  
 CC in the exemplification of the present invention

Sequence 575 BP; 124 A; 168 C; 165 G; 118 T; 0 U; 0 Other;

Query Match 18.5%; Score 495.8; DB 3; Length 575;  
 Best Local Similarity 94.8%; Pred. No. 6e-109;  
 Matches 532; Conservative 0; Mismatches 7; Indels 22; Gaps 1;

623 CAACATCGACGTCACCTCTTCCCAACAGCTCGGCTCTTCCAGAGAGCTTCTCCG 682

```
Db      10  CAAAGAGGCTTACAGACCTTTCCTCCACAGCTCGGCTCTTCCAGGAGCTTCTCTG 69
Qy      683 CAGGGGAGAGGCGTTCCTCAGTTGTGGGTCTGGGGGTGGTGCATCTCCCTAAAGAA 742
Db      70  CAGGGGAGAGGCGTTCCTCAGTTGTGGGTCTGGGGGTGGTGCATCTCCCTAAAGAA 129
Qy      743 GTGGTCCCTCTGTGGGCAAGTGAAGTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAA 802
Db      130 GTGGTCCCTCTGTGGGCAAGTGAAGTCTCCAGCTTTGGCTCAGCTCTCCAGACAGAA 189
Qy      803 AGACTGCGTCTTTTGTGGTGAACCCGAGGATCAGGAGATTGGAGCCCGTGAAGA 862
Db      190 AGACTGCGTCTTTTGTGGTGAACCCGAGGATCAGGAGATTGGAGCCCGTGAAGA 249
Qy      863 AGAAAATGAGAGG-----AGATGGGACCTTGAAGCTGAACGGG 900
Db      250 AGAAAATGAGAGGAGAAACATCCCTTCTCTGTAGATGGGGAAGCTTGAAGCTGAACGGG 309
Qy      901 CAGTTGTGTGTGGCACAACCGGTGAAATGCCCAAACCTTCCAGAGAGCTCAGAGCA 960
Db      310 CAGTTGTGTGTGGCACAACCGGTGAAATGCCCAAACCTTCCAGAGAGCTCAGAGCA 369
Qy      961 GCGGCTGGGAGGCGAGCAAGATGAGAGAGAGCGCTGACATGATCATCTGCCAGGACCTG 1020
Db      370 GCGGCTGGGAGGCGAGCAAGATGAGAGAGAGCGCTGACATGATCATCTGCCAGGACCTG 429
Qy      1021 CTGCACGACTCGGTGAGTTTGACAGCCCTGCAATGACACAGTTCTGGCGGGCTTGCTACTG 1080
Db      430 CTGCACGACTCGGTGAGTTTGACAGCCCTGCAATGACACAGTTCTGGCGGGCTTGCTACTG 489
Qy      1081 GGCTGGATGAGAGCGCTCGTCCCTGCTCCTACTGCGCGGTGCCGTGGAGCGGATCTGT 1140
Db      490 GGCTGGATGAGAGCGCTCGTCCCTGCTCCTACTGCGCGGTGCCGTGGAGCGGATCTGT 549
Qy      1141 AAAAAACACATCTCTCAACAAC 1161
Db      550 AAAAAACACATCTCTCAACAAC 570
```

Search completed: January 15, 2005, 00:57:44  
Job time : 1249 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 00:06:26 ; Search time 8052 Seconds

(without alignments) 12123.950 Million cell updates/sec

Title: US-10-048-046-1

Perfect score: 2679  
Sequence: 1 aagaattcgccagcgagcg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500.4	56.0	2297	3	HSN802304
2	1464.6	54.7	3146	3	AK077629
3	1366.2	51.0	2876	3	AK048930
4	1186.4	44.3	2412	3	CR625905
5	917.4	34.2	990	1	AL539798
6	911.2	34.0	1118	1	AL539797
7	822.2	30.7	876	5	BQ876252
8	774.8	28.9	921	5	BQ176559
9	765.4	28.6	870	5	BQ215731
10	758	28.3	992	5	BQ073354
11	757.6	27.6	910	5	CD243458
12	740.6	27.3	877	6	CD243458
13	727.4	27.2	840	4	BQ395714
14	726.8	27.1	953	5	BQ936072
15	714	26.7	938	5	BQ924736
16	708.4	26.4	824	4	BG831289
17	706.2	26.4	940	5	BX409790
18	700	26.1	909	5	BX409791
19	694.4	25.9	729	6	CA916737
20	694.4	25.9	1050	1	AL561570
21	674.8	25.2	897	1	BF971929
22	668.2	24.9	957	1	AL528043
23	661	24.7	940	5	BQ934867
24	652.4	24.4	759	4	BT69973

25	650.4	24.3	916	4	BM459647
26	648	24.2	694	7	CN428711
27	643.4	24.0	755	7	CN428709
28	639.6	23.9	666	2	BE407732
29	637.4	23.8	934	4	BF971838
30	637.2	23.8	1092	4	BM541730
31	628	23.4	687	1	AU138310
32	626.8	23.4	690	2	BE280165
33	626.2	23.4	938	5	BQ887242
34	624.8	23.3	1022	4	BM558904
35	624.8	23.3	1096	4	BG353562
36	617	23.0	722	1	AU135780
37	611	22.8	680	1	AU134937
38	602.6	22.5	636	2	BE407797
39	590.2	22.0	618	7	CN481126
40	588	21.9	635	4	BG481936
41	587.6	21.9	824	5	BQ572677
42	586.6	21.9	817	5	BM944161
43	583	21.8	824	1	AU132825
44	578.2	21.6	712	2	BE257482
45	567.8	21.2	650	2	AM957709

## ALIGNMENTS

RESULT 1  
HSN802304  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
CONSRM  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source  
1..2297  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="RZPD:DKFZp434N2420"  
/db\_xref="taxon:9606"  
/clone="DKFZp434N2420"  
/clone\_lib="434 (synonym: hhes3). Vector pSport1; host DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
/note="cell cycle checkpoint protein CHFR, N-terminus truncated"  
1..2297  
/gene="DKFZp434N2420"  
<1..921  
/gene="DKFZp434N2420"  
/codon\_start=1

2297 bp mRNA linear HTC 04-AUG-2004  
Homo sapiens mRNA; cDNA DKFZp434N2420 (from clone DKFZp434N2420).  
AL137561.1 GI:6808265  
HTC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 2297)  
Ostenwalder, B., Obermaier, B., Deutscher, S., Schaipe, A., Mewes, H.W., Weil, B., Amid, C., Oanger, A., Fob, G., Han, M. and Wiemann, S.  
The German cDNA Consortium  
Submitted (03-AUG-2004) MIPS, Ingolstaeder Landstr.1, D-85764 Neuherberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp434N2420) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp434N2420  
Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

```

/product="hypothetical protein"
/protein_id="CA870812.1"
/db_xref="GI:16808266"
/translation="VNEVYI QHDKRSREVDVOSMDANKITOMLQKPKRPSDDE
GSSEDLIELSNVDSSEISQPKVYRQCPEDRPOLADPBPQDAEERGPAPOLGSA
PSTVSULTVAYVQVCELOSALHTCTCCQPMFDRKAEEDPRVAPQOCVCLQPPC
HLTVGSTRIGEGGLAPFCEALGDKCLGVLNANNSTSEYDILKIKATLRTGTMKNMIL
EYVARNQGVFLSLDRAVYDVLQVLCGCGRSRRELYOYRONIPASELPAAVTSRPN
EYVARNQGVQKHAHNAKFNHILEQGRPN"

```

## ORIGIN

Query Match	Score	DB	Length
56.0%	1500.4	DB 3	2297

Matches 1501; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Matches 1501; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY	1165	GTGGAAAGCAATCTATCCAGATCAGACAAGAGTCGACGAAGAAAGATGTCAAAT	1122
Db	1	GTGGAAAGCAATCTATCCAGATCAGACAAGAGTCGACGAAGAAAGATGTCAAAGT	60
QY	1225	ATGGATGCGAGGAATAAATCACTCAAGACATGCTGCAGGCCCAAGTCAAGGGCTTTT	128
Db	61	ATGGATGCGAGGAATAAATCACTCAAGACATGCTGCAGGCCCAAGTCAAGGGCTTTT	120
QY	1285	TCTGATGAAGAAGGAGATTCAAGAGACTGTCTGAGCTTCAAGCGTTGACAGTAGTCC	1344
Db	121	TCTGATGAAGAAGGAGATTCAAGAGACTGTCTGAGCTTCAAGCGTTGACAGTAGTCC	180
QY	1345	TCAGACATTAACCGACCCATAGTCTGTGCTCCGGCAGTGTCTTGATGACAGAAAGCGAGG	140
Db	181	TCAGACATTAACCGACCCATAGTCTGTGCTCCGGCAGTGTCTTGATGACAGAAAGCGAGG	240
QY	1405	GGCGAGCCTCCCACTGCGCCAGACACCGAGGGCGAGCGAGAGCCCAACAGGCTCTGGG	146
Db	241	GGCGAGCCTCCCACTGCGCCAGACACCGAGGGCGAGCGAGAGCCCAACAGGCTCTGGG	300
QY	1465	GATGACCCCTTCCAGTCCCTCAGACTTGAACAGACATTCAGAGTTACGTGTCCCTCTG	152
Db	301	GATGACCCCTTCCAGTCCCTCAGACTTGAACAGACATTCAGAGTTACGTGTCCCTCTG	360
QY	1525	CAAGGAAGCCACGCGCTGTGCACTGTGCTTCCAGGCCATGCGCAGCGAGACCGAG	158
Db	361	CAAGGAAGCCACGCGCTGTGCACTGTGCTTCCAGGCCATGCGCAGCGAGACCGAG	420
QY	1585	CGCGAGCAGAACCGCGCTGTGGCCCTTCAAGAGTGTGGGTCTGCTTGCAGCTTTCTGC	1644
Db	421	CGCGAGCAGAACCGCGCTGTGGCCCTTCAAGAGTGTGGGTCTGCTTGCAGCTTTCTGC	480
QY	1645	CACCTGTACTGGGGGTGACACCGGACCGGCTGACGCTGGCGCCGTTTGTGAG	170
Db	481	CACCTGTACTGGGGGTGACACCGGACCGGCTGACGCTGGCGCCGTTTGTGAG	540
QY	1705	CTCAACCTGGGTGACAAAGTGTCTGGACGGCGTGTGAACAACAAACAGCTACAGTCAAC	176
Db	541	CTCAACCTGGGTGACAAAGTGTCTGGACGGCGTGTGAACAACAAACAGCTACAGTCAAC	600
QY	1765	ATCCGAAAGATTAACCTGGGCAACCGAGAGTTGACATGGAATAACATGTTGACCGAGAC	182
Db	601	ATCCGAAAGATTAACCTGGGCAACCGAGAGTTGACATGGAATAACATGTTGACCGAGAC	660
QY	1825	CTCGTGGCTTCCAGCGGGAAGTGTTCGTGCTGTGAATTACAGAGTCACGGAGACAC	1884
Db	661	CTCGTGGCTTCCAGCGGGAAGTGTTCGTGCTGTGAATTACAGAGTCACGGAGACAC	720
QY	1885	GTTCTGTGTTACTGTGTGGCTCTGGGCAAGCTTCCTGAGCTGACCTTATCAGTATGGGAG	1944
Db	721	GTTCTGTGTTACTGTGTGGCTCTGGGCAAGCTTCCTGAGCTGACCTTATCAGTATGGGAG	780
QY	1945	AACATTCCTGCTTCCGAGTTGCGAGTTCGAGTGCCTGATCCGCTCTGACTGTAATGGGAC	2004
Db	781	AACATTCCTGCTTCCGAGTTGCGAGTTCGAGTTCGAGTGCCTGATCCGCTCTGACTGTAATGGGAC	840
QY	2005	CGTAACTGCGCACTCAGGTGAAGCTCAACAACGCAATGAATTCATCATATCTGTAA	2064

Db	841	CGTAACTGCGCAGCTCAGGTGAAGAGCTCACACGCGCATGAATTCATCATATCTGTGAA	900
Qy	2065	CAGACAGAGTTCAAAAATAAGCATCCAGAGAGCCCTTGAGCAGCTTTCAGCACTGGAGGTG	212
Db	901	CAGACAGAGTTCAAAAATAAGCATCCAGAGAGCCCTTGAGCAGCTTTCAGCACTGGAGGTG	960
Qy	2125	AAGAGAGCGTGTTTTAAATAACAGAGACAAGACAAGTCAAGGTGTTTCACAGCCCCCTG	218
Db	961	AAGAGAGCGTGTTTTAAATAACAGAGACAAGACAAGTCAAGGTGTTTCACAGCCCCCTG	102
Qy	2185	AGGGAAGGGAACGACAGGCTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGCTTTT	224
Db	1021	AGGGAAGGGAACGACAGGCTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGCTTTT	108
Qy	2245	ACCGTCTGAGTGAACACCCCTCCCGAGAGCCCGGGGGCCGACGCCGCCCTCTGGTGAAC	230
Db	1081	ACCGTCTGAGTGAACACCCCTCCCGAGAGCCCGGGGGCCGACGCCGCCCTCTGGTGAAC	114
Qy	2305	GCTGGGGCAGGGCTCGTGGTGGCATCAGACGACAGACGAAAGCCCTTCTGTAAATGCGGC	236
Db	1141	GCTGGGGCAGGGCTCGTGGTGGCATCAGACGACAGACGAAAGCCCTTCTGTAAATGCGGC	120
Qy	2365	CGTCCCGCCGACAGGGGCGAGTTTGTCTTTTGTACATTTTCCGAAACATCAAGTTAAAGC	242
Db	1201	CGTCCCGCCGACAGGGGCGAGTTTGTCTTTTGTACATTTTCCGAAACATCAAGTTAAAGC	126
Qy	2425	AGAAGTCTGTTTTCAGGAAAAGTTTCAAGGAGAAAGGCAAGTTTATCAAAAACATTGTT	248
Db	1261	AGAAGTCTGTTTTCAGGAAAAGTTTCAAGGAGAAAGGCAAGTTTATCAAAAACATTGTT	132
Qy	2485	TCAGAGAGAGGAGCATTAAGTTTACAGCTTACAGGACGTACACAATATCTGCTGCGG	254
Db	1321	TCAGAGAGAGGAGCATTAAGTTTACAGCTTACAGGACGTACACAATATCTGCTGCGG	138
Qy	2545	AAAAACAAGCATTTTATCTAATTTTATTTAATGAGTGGTGGTCTATCTTCTAATTA	260
Db	1381	AAAAACAAGCATTTTATCTAATTTTATTTAATGAGTGGTGGTCTATCTTCTAATTA	144
Qy	2605	GATTTAAATGTACAAAACGTGAGACAAATATATTAATTTATTAATTTCAAAATTGACAA	266
Db	1441	GATTTAAATGTACAAAACGTGAGACAAATATATTAATTTATTAATTTCAAAATTGACAA	150
Qy	2665	AA 2666	
Db	1501	AA 1502	
RESULT 2			
AK077629			
LOCUS	3146 bp	mRNA	linear HTC 03-APR-2004
DEFINITION	Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730484M20 Product:similar to CELL CYCLE CHECKPOINT PROTEIN CHFR (Homo sapiens), full insert sequence.		
ACCESSION	AK077629		
VERSION	AK077629.1	GI:26346521	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		

PUBMED 11042159

REFERENCE 3

AUTHORS Shiba, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

TITLE RIKEN Integrated sequence analysis (RISA) system-384-format

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3146)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Konda, M., Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

PLEASE VISIT OUR WEB SITE FOR FURTHER DETAILS.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

LOCATION/Qualifiers

1. 3146

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM:DB:5730484M20"

/db\_xref="taxon:10090"

/clone="5730484M20"

/tissue\_type="whole body"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="8 days embryo"

89.. 2080

/note="unnamed protein product; putative similar to CELL CYCLE CHECKPOINT PROTEIN CHR (Homo sapiens) (SPTC109NR74, evidence: FASTA, 81.%ID, 100%length, match=1989)"

/codon\_start=1

/protein\_id="BAC36912.1"

/db\_xref="GI:26346522"

/translation="MEHGEDEPPPEPPPEMGLLTLGAEDBPOLLWKREWTIGRR  
GDDSPSNKLVSGDHKLTVDSEIAGEVTLSEPTNGVINKLQVKKQYPLSHGD1  
IYIVRRKRNPHNVAVIYESLSGKSLQDLSLEANKMPPHYTKQCSRGGDDPVP  
LISPAQCTLEBPQSTSTSDLPFASISSTREBELTSAGQKSSSSGSGNTSISPKR  
SSLVANGSLSSLPVFOKEASFSLSEKDEBELPARKKMGDESLTNQLVSGQ  
RGNAGTSEEDVADVPKDEAFSLTLCIDDLADCVSLQCHMTFAACVGMER  
SLICTCTCPERICNNHILNNVAYALIOHDKRSSEEDVSMARKITVDMLQPK  
VRRSPDEGSEEDLELSDVSESDISOPYIVGROCEPVRVAOSLPCVPESL  
CATLAGEABESTSILPAPDYMPILOGSHICTCCGOMPDRRAREERPAPOO  
CAVCIQPPCHLYMGTTRTGRCCLAPFELNLDKLDGVLNNNYEDILKNYATR  
GLTWVSVLTESLALQKRVFMLSIDRTITGNVLYCCCGKRSRELTQYRONI PASL  
PVTIVSRPDCVMGRNCRTOVNAHAKENHICEQTRFN"

polyA\_signal 3123.. 3128

polyA\_site 3146

/note="putative"

ORIGIN

Query Match 54.7%; Score 1464.6; DB 3; Length 3146;  
Best Local Similarity 74.5%; Pred. No. 0;  
Matches 2007; Conservative 0; Mismatches 634; Indels 52; Gaps 11;

18 CCGCAATGCTCTTGAACAGCGGCGCGCGG-----CAGCGGTTCCGGTTCCGCGCG 71

10 CCGAGATGCTGTGACAGCGGCGCGCGGCTGTAAGCCGCGTTCCGGATCCGCGCAG 69

72 GCGCGGATGTAATCCCGATGAGCGCGCCGAGAGAGGCAAGCACTGCCCGCCCGCA 131

70 GCGCGGAGTGAATCCCTATGAGCTACACAGGAGAGCGCGCGCGCGCGAGGA 129

132 GCGCGGAGAGCGGCTCTGCTGCTGCGGCGCGGAGAGGCGAGCGGCACTGCTCTGAG 191

130 ACCCTGGGAGAGGCTTCTTCTGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 189

192 GAAGCGGAGTGAACATCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251

190 GAAGCGGAGTGAACATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249

252 ACTGCTCTTGAAGATCACTGTAAGATTTGATGATGTAAGATTTGATGATGATGAT 311

250 ACTGCTCTTGAAGATCACTGTAAGATTTGATGATGATGATGATGATGATGATGAT 309

312 GGAAGATACACACACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371

310 GGAAGATACACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369

372 ATGCCCTTAAAGACTGCGGATGATCACTTGTGTGAAGAGAGAGAGAGAGAGAGAG 431

370 TTACCTTTAAAGAGAGAGAGAGAGATATCATTTATTTGATGAGAGAGAGAGAGAG 429

432 CAAGTGGATACCTCTATGAATCTTTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 491

430 CAATGGGATACCTCTATGAATCTTTAAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 489

492 TGAAGCTAACAG 551

490 TGAAGCTAACAG 543

552 GCGAGGAGCGGATCCCGGAGTCCCTCCGCTGCGCGCGGCACTGAGTGTGTTGAGGA 611

544 GCGAGGAGTATGATCCCGAGGTTCACTATGTCACCCAGGCTCAGAGATGTTGAGGA 603

612 ACCAGAGCATCAATGAGTCAAGAGCTCTTCCGACAGGCTGAGCTCTTCCAGGGA 671

604 ACCAGAGCATCAATGAGTCAAGAGCTCTTCCGACAGGCTGAGCTCTTCCAGGGA 663

672 GCC-----TTCTCGGAGGCGGAGAGGCTCTCCAGTTGGGTCTGGGGTGTGTG 725

664 GCGAGAGCTGAGCTCTGAGGCGGAGAGAGATCTTCTAGCTTGAAGAGAGAGAGAG 723

726 CATCTCCCTTAAGAGAGAGTGTCTCTGTGCAAGTGAAGTCTCAAGCTTTGCTTC 785

724 CATCTCCCTTAAGAGAGAGTGTCTCTGTGCAAGTGAAGTCTCAAGCTTTGCTTC 783



JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, S., Kawai, J., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE 5  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
PUBMED 1202876  
REFERENCE 6  
AUTHORS Aichi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kankawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akanita, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-42 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/  
location/Qualifiers  
1. 2876  
FEATURES  
Source  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM:DB:C230082M18"  
/db\_xref="taxon:10090"  
/clone="C230082M18"  
/tissue\_type="cerebellum"

misc\_feature  
/clone\_1ib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
1..2876  
/note="similar to CELL CYCLE CHECKPOINT PROTEIN CHFR [Homo sapiens] (SPTB) [Q9NRT4, evidence: FASTA, 81.7%id, 100%length, match=1989]"  
ORIGIN  
Query Match 51.0%; Score 1366.2; DB 3; Length 2876;  
Best Local Similarity 74.7%; Pred. No. 0;  
Matches 1862; Conservative 0; Mismatches 583; Indels 46; Gaps 10;  
214 CGGAGACGAGTGGCACCCTTCCTCCACCAATTAATGCTCTGAGATCACTGT 273  
390 CTGAAAAAGGCTGTGACCTCTCTTCCACCAATTAATGCTCTGAGATCACTGT 449  
274 AGAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 333  
450 AAACCTTACAGTGAATGAATATCTGGTGAAGTGAACCTGGAAGACCAATGGA 509  
334 ACAGGATTTAACAAGCTGAAGGTTGTTAAGAAGCAATGCTTACAGCTGGGAT 393  
510 ACAGTATCAATTAAGCTGCAAGTTGTTAAGAAGCAATGCTTACAGCTGGGAT 569  
394 GTCACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 453  
570 ATCATCTATTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 629  
454 TCTTTAAGTGAAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 513  
630 TCTTTAAGTGAAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 689  
514 TTTCAATGGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573  
690 TTTCAATGGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 743  
574 CCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633  
744 CCAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803  
634 TGAAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 687  
804 TGAAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863  
688 CAGAGCGCTCTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747  
864 CAAAGCATTTCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
748 CCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807  
924 TCACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980  
808 GGTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867  
981 GCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040  
868 ATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 926  
1041 GATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1100  
927 AATGCGCCAAACCTGCAAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 986  
1101 AATGCGCCAAACCTGCAAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1160  
987 GAGAGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1046  
1161 GAGAGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220  
1047 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1106  
1221 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1280  
1107 TCTACTGCGCGCTGTCCCGTGAAGCGGATGATGATGATGATGATGATGATGATGATGAT 1166





## ORIGIN

Query Match 44.3%; Score 1186.4; DB 3; Length 2412;  
 Best Local Similarity 95.4%; Pred. No. 8.4e-296;  
 Matches 1257; Conservative 0; Mismatches 51; Indels 10; Gaps 3;

1354 AGCCAGCCATACGTCGTGTGCGGAGTGTCTGAGTACAGAGGCGGCGAGCCCT 1413  
 530 AGCCAGCCATACGTCGTGTGCGGAGTGTCTGAGTACAGAGGCGGCGAGCCCT 589  
 1414 CCCCATGTCGACACCCGAGGCGAGCCAGAGGCCCAAGGCCCTGCGGAGTGCACC 1473  
 590 CCCCATGTCGACACCCGAGGCGAGCCAGAGGCCCAAGGCCCTGCGGAGTGCACC 649  
 1474 TCCAGCTCCGTGAGCTGACAGAGATTCAGAGTTAGTGTGCTCTGCAAGAGAC 1533  
 650 TCCAGCTCCGTGAGCTGACAGAGATTCAGAGTTAGTGTGCTCTGCAAGAGAC 706  
 1534 CAGGCGCTGTGCACTGTGCTTCCAGCCCATGCGGAGCGGAGAGCGGAGCGAG 1593  
 707 CAGGCGCTGTGCACTGTGCTTCCAGCCCATGCGGAGCGGAGAGCGGAGCGAG 766  
 1594 GACCCGCTGTGCGCCCTCAGCAGTGTGCGTCTGCTGAGCCCTTCTGCCACTGTAC 1653  
 767 GACCCGCTGTGCGCCCTCAGCAGTGTGCGTCTGCTGAGCCCTTCTGCCACTGTAC 826  
 1654 TGGGGCTGCACTGCGACCGGCTGCTACGGTGTGCTGCGCCCGTTTGTGAGTCACTG 1713  
 827 TGGGGCTGCACTGCGACCGGCTGCTACGGTGTGCTGCGCCCGTTTGTGAGTCACTG 886  
 1714 GGTGACAGTGTGCACTGCGGCTGCGACCAACAGAGTACAGAGTCACTGCA-- 1771  
 887 AAAAGAAATATGTTATCTCTGAGTGTGCGGACAGTGTCTTCTGCGACATATTAA 946  
 1772 ---AGAAATTAATCTGCGACAGAGTGTGACATGTAAGAAACATGTTGACCGAGGCTCG 1828  
 947 AACAGAAATTAATCTGCGACAGAGTGTGACATGTAAGAAACATGTTGACCGAGGCTCG 1006  
 1829 TGGCTCTCAGCGGAGGAGTGTGTCTGTCTGATTAACAGAGTCACTGCGGAGACCTGTC 1888  
 1007 TGGCTCTCAGCGGAGGAGTGTGTCTGTCTGATTAACAGAGTCACTGCGGAGACCTGTC 1066  
 1889 TGTGTTACTGCTGTGCGCTGCGGAGCTCCGTAAGCTGACATGTAATCTGTAAGAGAA 1948  
 1067 TGTGTTACTGCTGTGCGCTGCGGAGCTCCGTAAGCTGACATGTAATCTGTAAGAGAA 1126  
 1949 TTTCTGCTTCCAGAGTGTGCGGAGTGTGTCTGTCTGATTAACAGAGTCACTGCGGAG 2008  
 1127 TTTCTGCTTCCAGAGTGTGCGGAGTGTGTCTGTCTGATTAACAGAGTCACTGCGGAG 1186  
 2009 ACTGCGGCTCAGGTGAAAGCTCAGCAGGCTGTAATCAATCAATCTGTGAACAGA 2068  
 1187 ACTGCGGCTCAGGTGAAAGCTCAGCAGGCTGTAATCAATCAATCTGTGAACAGA 1246  
 2069 CAAAGTTCAAAACCTAAGATCCAGAGGCGCTGAGAGCTTTCAGACCTGAGAGTAA 2128  
 1247 CAAAGTTCAAAACCTAAGATCCAGAGGCGCTGAGAGCTTTCAGACCTGAGAGTAA 1306  
 2129 GAGCGTGTGTTTAAATTAAGAGACAGACGTCAGAGTGTGTTTTCAGAGCCCTGAGAG 2188  
 1307 GAGCGTGTGTTTAAATTAAGAGACAGACGTCAGAGTGTGTTTTCAGAGCCCTGAGAG 1366  
 2189 AAGGAGCGAGAGGTCTCCAGAGAGTGTCTGAGAGTGTGTTTTCAGAGCTTTTAAACC 2248  
 1367 AAGGAGCGAGAGGTCTCCAGAGAGTGTCTGAGAGTGTGTTTTCAGAGCTTTTAAACC 1424  
 2249 TGTGAGTGAAGCCCTCCAGAGGCGCGGAGGCGGAGCGGCGCTCTGAGAGGCGTG 2308  
 1425 TGTGAGTGAAGCCCTCCAGAGGCGCGGAGGCGGAGCGGCGCTCTGAGAGGCGTG 1484  
 2309 GCGAGGCTCTGTGTGTGATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2368  
 1485 GCGAGGCTCTGTGTGTGATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1544

QY 2369 CCGCCGAGAGGCGAGTGTGCTTGTGATCTTTCCGAAACTACAGTTAAAGCAGA 2428  
 DB 1545 CCGCCGAGAGGCGAGTGTGCTTGTGATCTTTCCGAAACTACAGTTAAAGCAGA 1604  
 QY 2429 GTCTGTTTTCAGAGAAAGTTTCAAGGAGAGGCAAGTTTACAAAACATGTTTCAG 2488  
 DB 1605 GTCTGTTTTCAGAGAAAGTTTCAAGGAGAGGCAAGTTTACAAAACATGTTTCAG 1664  
 QY 2489 GAGAGGAGCAATTAATTAACAGCTTACAGAGAGTACACATATCTGCTGCGGAGAA 2548  
 DB 1665 GAGAGGAGCAATTAATTAACAGCTTACAGAGAGTACACATATCTGCTGCGGAGAA 1724  
 QY 2549 CCAAGCATTTATATCTTATTTTATTTATAGTGTGCTGCTATCTTATTAAGATT 2608  
 DB 1725 CCAAGCATTTATATCTTATTTTATTTATAGTGTGCTGCTATCTTATTAAGATT 1784  
 QY 2609 TAAATGTCAAAACTGTAGCACAATAATATATATTAATTAACAAATTGACAAA 2666  
 DB 1785 TAAATGTCAAAACTGTAGCACAATAATATATATTAATTAACAAATTGACAAA 1842

RESULT 5  
 AL539798 990 bp mRNA linear EST 24-MAR-2004  
 LOCUS AL539798 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
 DEFINITION CS0DF031YA23 5-PRIME, mRNA sequence.  
 ACCESSION AL539798  
 VERSION AL539798.3 GI:45715436  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31264361.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: sequef@genoscope.cns.fr; Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6792.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?c=CS0DF031YA120P1ec=6792.r.  
 Location/Qualifiers

FEATURES  
 source

1..990  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF031YA23"  
 /tissue\_type="FETAL BRAIN"  
 /dev\_stage="fetal"  
 /clone\_1ib="Homo sapiens FETAL BRAIN"  
 /note="Torgan: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Query Match 34.2%; Score 917.4; DB 1; Length 990;  
 Best Local Similarity 99.2%; Pred. No. 3.6e-226;  
 Matches 937; Conservative 5; Mismatches 1; Indels 2; Gaps 2;

61 GTTCCGCGGCGGCGGAGTGTGATCCCGATGAGCGGCGGAGAGAGAGAGAGAGAGAG 120

```

Db      20 GTTCCGCGCGGCGGGAATGTAATCCGATGAGCGGCCGAGAGGCAAGCATGCG 79
Qy      121 CCGCGCGCGGAGCCCTGAGGAGCGGCTCTGCGTCTGGGCGGAGAGGCGGAGCGCAC 180
Db      80 CCGCGCGCGGAGCCCTGAGGAGCGGCTCTGCGTCTGGGCGGAGAGGCGGAGCGCAC 139
Qy      181 GTTCTCTTGAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db      140 GTTCTCTTGAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 199
Qy      241 CCAGCAATAAATGCTGCTCTGGAAGATCACTGTAATGTAATGTAATGTAATGTAAT 300
Db      200 CCAGCAATAAATGCTGCTCTGGAAGATCACTGTAATGTAATGTAATGTAATGTAAT 259
Qy      301 CAGGTGACACTGGAAGATACAGCAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 360
Db      260 CAGGTGACACTGGAAGATACAGCAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 319
Qy      361 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      320 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
Qy      421 GAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db      380 GAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439
Qy      481 CAGAGATCTTTGAAAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      440 CAGAGATCTTTGAAAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 499
Qy      541 GCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db      500 GCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559
Qy      601 TGCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      560 TGCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 619
Qy      661 TCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      620 TCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 679
Qy      721 GGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      680 GGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
Qy      781 GCTCAGCTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db      740 GCTCAGCTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 799
Qy      841 GAGGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db      800 GAGGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858
Qy      901 CAGTGTGTTGTCGCAACACCGGTAGAAATGCGCAACCGGTAGAAATGCGCAACCGGT 960
Db      859 CAGTGTGTTGTCGCAACACCGGTAGAAATGCGCAACCGGTAGAAATGCGCAACCGGT 918
Qy      961 GCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
Db      919 GCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962

```

```

RESULT 6
AL539797/c 1118 bp mRNA linear EST 24-MAR-2004
LOCUS      AL539797 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF031YA23 3-PRIME, mRNA sequence.
ACCESSION  AL539797
VERSION     AL539797.3 GI:45715435
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens

```

```

REFERENCE 1 (bases 1 to 1118)
AUTHORS   Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    On Feb 15, 2001 this sequence version replaced gi:31264360.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 9106 Evry cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6792.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=CS0DF031YA12NP1&c=6792.r.

FEATURES
source
1..1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF031YA23"
/lisue_type="FETAL BRAIN"
/dev_stage="Fetal"
/clone_id="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 34.0%; Score 911.2; DB 1; Length 1118;
Best Local Similarity 91.3%; Pred. No. 1.5e-224;
Matches 973; Conservative 40; Mismatches 47; Indels 6; Gaps 5;

Qy      1580 CGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639
Db      1060 CCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1001
Qy      1640 TCTGCACTCTGTAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1699
Db      1000 TCTGCACTCTGTAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944
Qy      1700 GTGAGCTCAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1759
Db      943 GAGGCTCAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
Qy      1760 CAGACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
Db      883 CAGACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 824
Qy      1820 AGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1879
Db      823 AGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 764
Qy      1880 ACAACCTTCTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1939
Db      763 ACAACCTTCTGTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 704
Qy      1940 GGCAGAACATCTCTGCTTCCGAGTTGCGAGTGGCGGTAATCCGCTGAGCTGTAAT 1999
Db      703 GGCAGAACATCTCTGCTTCCGAGTTGCGAGTGGCGGTAATCCGCTGAGCTGTAAT 644
Qy      2000 GGGGCGTAATCTCGGAGCTCAGGTGAAGAGTCAACAGCGCATGAATTAATATATCT 2059
Db      643 GGGGCGTAATCTCGGAGCTCAGGTGAAGAGTCAACAGCGCATGAATTAATATATCT 584
Qy      2060 GTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2119

```

Db 583 GTGAACAGACAAGGTTCAAAAATAAGCATCCAGAGCCCTTGAGCAGCTTTCAGACACTG 524  
 Qy 2120 AGGTAAAGAGACGCTCTTTTAAATAAGAGACAGCAACGTCACAGCTTTTCAAGCC 2179  
 Db 523 AGGTAAAGAGACGCTCTTTTAAATAAGAGAGCAAGCAAGCTCAGCTTTTCAAGCC 464  
 Qy 2180 CCTGAG 2239  
 Db 463 CCTGAG 404  
 Qy 2240 TTTTAACTCTGAGTGAAGACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2299  
 Db 403 --TTTAACTCTGAGTGAAGACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346  
 Qy 2300 TGAAGGCTTGGGAG 2359  
 Db 345 TGAAGGCTTGGGAG 286  
 Qy 2360 GGGGCGCTCCCGGAG 2419  
 Db 285 GGGGCGCTCCCGGAG 226  
 Qy 2420 AAAGCAGAGAGTCTGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2479  
 Db 225 AAAGCAGAGAGTCTGTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166  
 Qy 2480 TTGTTTTCAG 2539  
 Db 165 TTGTTTTCAG 106  
 Qy 2540 CTGGAG 2599  
 Db 105 CTGGAG 46  
 Qy 2600 AATAAGATTAAATGTCACAACTGAGCACAATAATAATTA 2645  
 Db 45 AATAAGATTAAATGTCACAACTGAGCACAATAATAATTA 1  
 RESULT 7  
 BQ876252 876 bp mRNA linear EST 16-AUG-2002  
 LOCUS BQ876252  
 DEFINITION AGENCOURT 8681666 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6376937  
 ACCESSION BQ876252  
 VERSION BQ876252.1 GI:22268258  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 876)  
 NIH-MGC http://mgi.mgi.nhl.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgsabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L10C2559 row: 0 column: 18  
 High quality sequence stop: 678.  
 Location/Qualifiers  
 1. 876  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6376937"  
 /issue\_type="normal pigmented retinal epithelium"

/lab host="DH10B (phage-resistant)"  
 /clone lib="NIH\_MGC\_43"  
 /note="Organ: eye; Vector: pOTB1; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 30.7%; Score 822.2; DB 5; Length 876;  
 Best Local Similarity 98.3%; Pred. No. 1.7e-201;  
 Matches 863; Conservative 0; Mismatches 9; Indels 6; Gaps 3;

1165 GTGAAGACATACCTCATCCAGATCCAGACAGAGAGTCCGAGAGAGAGAGAGAGAG 1224  
 Db 1 GTGAAGACATACCTCATCCAGATCCAGACAGAGAGTCCGAGAGAGAGAGAGAGAG 60  
 Qy 1225 ATGATGCCAGAAATTAATCACTCAAGACATGCTGACGCCAAGTCAAGCGCTTTT 1284  
 Db 61 ATGATGCCAGAAATTAATCACTCAAGACATGCTGACGCCAAGTCAAGCGCTTTT 120  
 Qy 1285 TCTGATGAAGAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1344  
 Db 121 TCTGATGAAGAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 Qy 1345 TCAGACATTAGCAGACCATACGTCGTGTCGCGGAGAGTCTGAGTACAGAGAGAGAG 1404  
 Db 181 TCAGACATTAGCAGACCATACGTCGTGTCGCGGAGAGTCTGAGTACAGAGAGAGAG 240  
 Qy 1405 GCGCAGCTCCCACTGCCAGACACCCGAGGCGAGCCAGAGAGAGAGAGAGAGAGAG 1464  
 Db 241 GCGCAGCTCCCACTGCCAGACACCCGAGGCGAGCCAGAGAGAGAGAGAGAGAGAG 300  
 Qy 1465 GATGACACCTCCAGAGTCCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524  
 Db 301 GATGACACCTCCAGAGTCCGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 357  
 Qy 1525 CAAG 1584  
 Db 358 CAAG 417  
 Qy 1585 GCGAG 1644  
 Db 418 GCGAG 477  
 Qy 1645 CACCTGATCTGGGAGTGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1704  
 Db 478 CACCTGATCTGGGAGTGAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537  
 Qy 1705 CTCAACTGAGTGAAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1764  
 Db 538 CTCAACTGAGTGAAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597  
 Qy 1765 ATCTGAAGAAATTAATCTGCAACAGAGAGTTCAGATGAGAGAGAGAGAGAGAG 1824  
 Db 598 ATCTGAAGAAATTAATCTGCAACAGAGAGTTCAGATGAGAGAGAGAGAGAGAGAG 657  
 Qy 1825 CTCGAGGCTTCACAGGAG 1882  
 Db 658 CTCGAGGCTTCACAGGAG 717  
 Qy 1883 CCGTCTGTTTAACTGCTGTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1942  
 Db 718 CCGTCTGTTTAACTGCTGTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777  
 Qy 1943 AGAATATTCCTGCTTCCAGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2002  
 Db 778 AGAATATTCCTGCTTCCAGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837  
 Qy 2003 -GCCGTAATCTGCCGACATCAGGTGAAGACTACCAAGC 2039





[illegible]

Db	4	AGGAAGCCACGCCCTGTGTGACCTGTGCTTTCAGGCCCATGCCCCAGCCGAGACGCGAGCG	63
QY	1587	CGAGCAGGACCCGCGCTGTGCGCCCTTCAGCAGTGTGCGGTCTTGCCCTGACGCCCTTTCGCA	1646
Db	64	CGAGCAGGACCCGCGCTGTGCGCCCTTCAGCAGTGTGCGGTCTTGCCCTGACGCCCTTTCGCA	123
QY	1647	CCTGTACTGGGGGTGCAACCCGAGCCGGTGTCTAGGGCTGCCGGCCCGTTTGTGACT	1706
Db	124	CTGTACTGGGGGTGCAACCCGAGCCGGTGTCTAGGGCTGCCGGCCCGTTTGTGACT	183
QY	1707	CAACCTGGGTGACAGTGTCTGGAACGCGGTCTGAAACAACAGCTACGAGTACAGAT	1766
Db	184	CAACCTGGGTGACAGTGTCTGGAACGCGGTCTGAAACAACAGCTACGAGTACAGAT	243
QY	1767	CCTGAAGAAATTAACCTGCGCAACCAAGGTTTGAATGAAAAATGTTGACCGAGAGCTT	1822
Db	244	CCTGAAGAAATTAACCTGCGCAACCAAGGTTTGAATGAAAAATGTTGACCGAGAGCTT	303
QY	1827	CGTGGCTCTCCAGCGGGAGTGTTCGTCTGTGATTAACAAGTCAAGGAGACACCGT	1886
Db	304	CGTGGCTCTCCAGCGGGAGTGTTCGTCTGTGATTAACAAGTCAAGGAGACACCGT	363
QY	1947	CATTCCGCTCCGAGTTGCCAGTGGCGCGTAAATCCCGTCTGACTGTCTGCGGGCG	2006
Db	424	CATTCCGCTCCGAGTTGCCAGTGGCGCGTAAATCCCGTCTGACTGTCTGCGGGCG	483
QY	2007	TAACTGCGCAGCTCAGGTGAAGTCTCAACAAGCATGAATTCATCATATCTGTGAACA	2066
Db	484	TAACTGCGCAGCTCAGGTGAAGTCTCAACAAGCATGAATTCATCATATCTGTGAACA	543
QY	2067	GACAAAGTTCAAAAACCTAAGATTCACAGAGCCCTGAGCAGCTTTCAGCATCTGAGGTGA	2126
Db	544	GACAAAGTTCAAAAACCTAAGATTCACAGAGCCCTGAGCAGCTTTCAGCATCTGAGGTGA	603
QY	2127	GAGAGCGTGTTTTAAATATACAGAGACAGACGTCAAGGTTTTCACAGAGCCCTGTAG	2186
Db	604	GAGAGCGTGTTTTAAATATACAGAGACAGACGTCAAGGTTTTCACAGAGCCCTGTAG	663
QY	2187	GGAAGGGAACGAGGCTCTCCGACAGGTGCTCTGGGGTGAATCTTCTGTGAGCTTTTAC	2246
Db	664	GGAAGGGAACGAGGCTCTCCGACAGGTGCTCTGGGGTGAATCTTCTGTGAGCTTTTAC	723
QY	2247	CCTCTGAGTGAAGACCTCCCGCAGAGCCCGGGGGCGGAGCGCCCTCCGTGAGAGCG	2306
Db	724	CCTCTGAGTGAAGACCTCCCGCAGAGCCCGGGGGCGGAGCGCCCTCCGTGAGAGCG	783
QY	2307	TGGGCAAGGGCTCTGTGT	2323
Db	784	GCGCTTGGGCAAGGGCT	800
RESULT 12			
BUI48471			
LOCUS	BUI48471	910 bp	mRNA linear EST 03-SEP-2002
DEFINITION	AGNCOURT 8675647 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380102		
	5', mRNA sequence.		
ACCESSION	BUI48471		
VERSION	BUI48471.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 910)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strauberg, Ph.D.		



Email: c9apbs-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
 Plate: LNCM2568 row: c column: 15  
 High quality sequence stop: 620.  
 Location/Qualifiers

## FEATURES

source

1. 910  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6380102"  
 /issue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 40"  
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI, cDNA made by oligo-dt priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

## ORIGIN

## Query Match

Best Local Similarity 91.1%; Pred. No. 2.5e-180;  
 Matches 858; Conservative 0; Mismatches 40; Indels 44; Gaps 5;

QY 12 AGGAGGCGCATGTTCTTTGACAGCGGCGCGCCGACCGGTTCCGGTTGCGCGCG 71  
 DB 5 AGGAGGCGCATGTTCTTTGACAGCGGCGCGCGCCGACCGGTTCCGGTTGCGCGCG 64  
 QY 72 GCGCGGAGTGTGATCCGATGGAGGCGCGCGAGGAGGAGGAGTCCGCGCGCGCGA 131  
 DB 65 GCGCGGAGTGTGATCCGATGGAGGCGCGCGAGGAGGAGGAGTCCGCGCGCGCGA 124  
 QY 132 GCGCTGGGAGCGGCTCTGCGTGTGGCGCGAGAGGCGCGAGCGCCGACGTCCTCTGAG 191  
 DB 125 GCGCTGGGAGCGGCTCTGCGTGTGGCGCGAGAGGCGCGAGCGCCGACGTCCTCTGAG 184  
 QY 132 GCGCTGGGAGCGGCTCTGCGTGTGGCGCGAGAGGCGCGAGCGCCGACGTCCTCTGAG 191  
 DB 125 GCGCTGGGAGCGGCTCTGCGTGTGGCGCGAGAGGCGCGAGCGCCGACGTCCTCTGAG 184  
 QY 185 GAAAGCGGAGTGGACCATCGGCGGAGAGAGAGGTTCCGACCTTCTCCCGACAAATTA 244  
 DB 252 ACTGGTCTGAGAGTCACTGATGATTTGATGAGTGAATAATCGATGAGTCACT 311  
 DB 245 ACTGGTCTGAGAGTCACTGATGATTTGATGAGTGAATAATCGATGAGTCACT 304  
 QY 312 GGAAGATACCAAGACCAAGTGAACAGTGAATGAAGTGTGAAGAGAGAC 371  
 DB 305 GGAAGATACCAAGACCAAGTGAACAGTGAATGAAGTGTGAAGAGAGAC 364  
 QY 372 ATGCCCTTTACAGACTGGGAGTGTCTACTTGTGTGACAGAGAGATGAACCGGAGAC 431  
 DB 365 ATGCCCTTTACAGACTGGGAGTGTCTACTTGTGTGACAGAGAGATGAACCGGAGAC 424  
 QY 432 CAACGCGGATACCTTATGATCTTTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 491  
 DB 425 CAACGCGGATACCTTATGATCTTTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 484  
 QY 492 TGAAGCTAACAGAGAAATGTCTTCATGGAGCAAGAGATACCTCAGTGCAGTGCAGG 551  
 DB 485 T-----GATACCTCAGTGCAGTGCAGG 508  
 QY 552 GCGAGGCGCGATCCCGGAGTCCCTCGTGTGCGCCGACATCAGGTGTGCTTTGAGAG 611  
 DB 509 GCGAGGCGCGATCCCGGAGTCCCTCGTGTGCGCCGACATCAGGTGTGCTTTGAGAG 568

QY 612 ACCACAGCATCAATCATGACGTCAGACCTCTTCCACAGGCTCGGCTCTTCCACGGA 671  
 DB 569 ACCACAGCATCAATCATGACGTCAGACCTCTTCCACAGGCTCGGCTCTTCCACGGA 628  
 QY 672 GCTTCTCTGACAGGCGGAGAGCGTTCTTCAGTTGTGGGTCTGGGGGTGTGGCATCTC 731  
 DB 629 GCTTCTCTGACAGGCGGAGAGCGTTCTTCAGTTGTGGGTCTGGGGGTGTGGCATCTC 688  
 QY 732 CCTTAAGAGAGTGTCTCTGTGCGAAGTATGAAGTCTCAGCTTTGCCCTCACTCT 791  
 DB 689 CCTTAAGAGAGTGTCTCTGTGCGAAGTATGAAGTCTCAGCTTTGCCCTCACTCT 748  
 QY 792 CCACAGAGAGAGACT-GGCTCTTTCTGCTGTGAGACCCAGATCAGAGAGATTGG 850  
 DB 749 CCACAGAGAGAGACTGGGTCTTTCTCGGTGAGACCCAGATCAGAGAGATTGG 808  
 QY 851 AGCCCTGTAAGAGAGAGATGA---GAGAGATGGGAGCTTTG-ACCTGAACGGGCACT 904  
 DB 809 GGAAGCCGTAAGAGAGAGATGA---GAGAGATGGGAGCTTTG-ACCTGAACGGGCACT 868  
 QY 905 TGTGTGTCGACAGACC-GGCTAGAAATGCCCAACCTGTCAC 945  
 DB 869 GTGTGTCCCAACAGCGGGTGAATGCCCAACAGGCGCC 910

## RESULT 13

BG395714

LOCUS

DEFINITION

602458365P1 NIH\_MGC\_16 Homo sapiens cDNA clone IMAGE:458063 5',  
 mRNA sequence.

ACCESSION BG395714  
 VERSION BG395714.1 GI:13289162

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLES

JOURNAL

COMMENT

NIH-MGC <http://mgi.mci.nih.gov/>.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LNCM1300 row: a column: 18  
 High quality sequence stop: 775.  
 Location/Qualifiers

## FEATURES

source

1. 840  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4580633"  
 /issue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 16"  
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

## ORIGIN

## Query Match

Best Local Similarity 95.8%; Pred. No. 6.4e-177;  
 Matches 779; Conservative 0; Mismatches 31; Indels 3; Gaps 3;

QY 1294 GAAGGAGTTGAGAGGAGCTGCTGAGAGCTGTGACAGCTTTGACAGTGAAGTCTTCAAGCATT 1353  
 DB 2 GAAGGAGTTGAGAGGAGCTGCTGAGAGCTGTGACAGCTTTGACAGTGAAGTCTTCAAGCATT 61  
 QY 1354 AGCCAGCCTAGAGTGTGTGCGGGAGTGTCTTGAAGTGAAGAGAGGCGGCGAGCCT 1413  
 DB 62 AGCCAGCCTAGAGTGTGTGCGGGAGTGTCTTGAAGTGAAGAGAGGCGGCGAGCCT 121  
 QY 1414 CCCCATGCGCCAGCAGCCGAGGGCGAGCCAGAGAGCCCAAGAGCCCTGAGGGATGACACC 1473  
 DB 122 CCCCATGCGCCAGCAGCCGAGGGCGAGCCAGAGAGCCCAAGAGCCCTGAGGGATGACACC 181  
 QY 1474 TCCAGCTCCGTGAGCTGAGCAGACAGAGTTCAGAGATTAGTGTGCCCTTGTCAAGAGAC 1533  
 DB 182 TCCAGCTCCGTGAGCTGAGCAGACAGAGTTCAGAGATTAGTGTGCCCTTGTCAAGAGAC 241  
 QY 1534 CAGCCCTGTGACCTGCTGCTTCCAGGCCATGCCGAGCCGAGAGAGCGGAGCGAGCAG 1593  
 DB 242 CAGCCCTGTGACCTGCTGCTTCCAGGCCATGCCGAGCCGAGAGAGCGGAGCGAGCAG 301  
 QY 1594 GACCCGCTGTGCGCCCTCAGAGAGTGTGCGGTCTGCTGAGCCTTCTGCACTGTAC 1653  
 DB 302 GACCCGCTGTGCGCCCTCAGAGAGTGTGCGGTCTGCTGAGCCTTCTGCACTGTAC 361  
 QY 1654 TGGGGGCTGACCCCGAGCCGCGCTGACAGCCGCTGAGCCGCTTTTGTGAGTCAACTG 1713  
 DB 362 TGGGGGCTGACCCCGAGCCGCGCTGACAGCCGCTGAGCCGCTTTTGTGAGTCAACTG 421  
 QY 1714 GGTGACAAAGTGTCTGACCGCGCTGCTGACAAACAGCTACAGAGTCAAGATCCTGAAG 1773  
 DB 422 GGTGACAAAGTGTCTGACCGCGCTGCTGACAAACAGCTACAGAGTCAAGATCCTGAAG 481  
 QY 1774 AATTACTGTGCAACCAAGAGTTTGAACATGGAAGAAATCATTTGACCGAGAGCTGTGCT 1833  
 DB 482 AATTACTGTGCAACCAAGAGTTTGAACATGGAAGAAATCATTTGACCGAGAGCTGTGCT 541  
 QY 1834 CTCGACCGGGAGTGTCTGCTGTCTGATTAAGAGTCAAGGAGAGACACCGTCTGTGT 1893  
 DB 542 CTCGACCGGGAGTGTCTGCTGTCTGATTAAGAGTCAAGGAGAGACACCGTCTGTGT 601  
 QY 1894 TACTGCTGTGCGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGACAGAACTTCT 1953  
 DB 602 TACTGCTGTGCGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGACAGAACTTCT 661  
 QY 1954 GTTTCCGAGTTGCCAGTGGCGCTTAATCCCTCTGACTGCTACTGCGGCGCTTAATCTGC 2013  
 DB 662 GTTTCCGAGTTGCCAGTGGCGCTTAATCCCTCTGACTGCTACTGCGGCGCTTAATCTGC 721  
 QY 2014 CGCACTCA-GGTGAAGCTC-ACCAAGCCATGAATTAATCATATCTGT-GAAGACAGA 2070  
 DB 722 GCACTCAAGGTTGAAGAGCTCAACCAAGCCATGAATTAATCATATCTGTGAGACAGAG 781  
 QY 2071 AGTTCAAAAATAAGATCCAGAGGCGCTTGAAG 2103  
 DB 782 AGTTCAAAAATAAGATCCAGAGGCGCTTGAAG 814  
 RESULT 14  
 B0936072 953 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGENCOURT 8779684 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:5371968  
 DEFINITION 5', mRNA sequence.  
 ACCESSION B0936072  
 VERSION B0936072.1 GI:22351455  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 953)  
 AUTHORS NIH-MGC <http://nigc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DRP/Gazdar  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
<http://image.llnl.gov>  
 Plate: ILCM2546 row: P column: 17  
 High quality sequence stop: 630.  
 FEATURES  
 source  
 1..953  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5371968"  
 /tissue\_type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_18"  
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN  
 Query Match 27.1%; Score 726.8; DB 5; Length 953;  
 Best Local Similarity 93.0%; Pred. No. 9.5e-177;  
 Matches 842; Conservative 0; Mismatches 18; Indels 45; Gaps 6;  
 45 GCGCAGCGGTTCCGGGTTGCGGCGGGGCGGAGATGATCCGAGTGAAGCGGCGCGA 104  
 DB 1 GCGCAGCGGTTCCGGGTTGCGGCGGGGCGGAGATGATCCGAGTGAAGCGGCGCGA 60  
 QY 105 GAAAGGCAAGCAGTGCAGCGCGCGCGAGCCCTGAGGAGCGGCTCTGCTGCGCGGCA 164  
 DB 61 GAAAGGCAAGCAGTGCAGCGCGCGCGAGCCCTGAGGAGCGGCTCTGCTGCGCGGCA 120  
 QY 165 GAAAGGCGAGCGCGAGCTCTCTGAGAGAGCGGAGTGAACCATCGGCGGAGACGAG 224  
 DB 121 GAAAGGCGAGCGCGAGCTCTCTGAGAGAGCGGAGTGAACCATCGGCGGAGACGAG 180  
 QY 225 TTGGCACTTCTCTCCCGCAATAAATGCTCTGAGATCACTGTAATTTGTAGT 284  
 DB 181 TTGGCACTTCTCTCCCGCAATAAATGCTCTGAGATCACTGTAATTTGTAGT 240  
 QY 285 GATGAATAAATCAGTGAAGTGAACATGAAATACAGACAGAGTGAACAGTATTA 344  
 DB 241 GATGAATAAATCAGTGAAGTGAACATGAAATACAGACAGAGTGAACAGTATTA 300  
 QY 345 CAAGCTGAAGTTGTTAAGAGAGAGATGCCCTTTACATCGGGATGTCATCTT 404  
 DB 301 CAAGCTGAAGTTGTTAAGAGAGAGATGCCCTTTACATCGGGATGTCATCTT 360  
 QY 405 GGTGTACAGAGAGATGAACCGAGAACAGAGTGGCATACCTCATGAATCTTTAATGA 464  
 DB 361 GGTGTACAGAGAGATGAACCGAGAACAGAGTGGCATACCTCATGAATCTTTAATGA 420  
 QY 465 AAAGCAAGCATGACACAGAAATCTTTGAAGCTTAACAGAGAAATGTTCATGGAC 524  
 DB 421 AAAGCAAGCATGACACAGAAATCTTT----- 448  
 QY 525 CAAGATACCTCAGGTGACGTGACGAGGAGGAGGCGGATCCCGGCTCCCTCGTGTG 584  
 DB 449 ----GATACCTCAGGTGACGTGACGAGGAGGAGGCGGATCCCGGCTCCCTCGTGTG 504  
 QY 585 GCCCGCACTCAGGTGCTTTGAGAGAACCAAGCATCAATCAAGTCAAGCTCTT 644

Db 505 GCCCGCACTCAGTGTGCTTTGAGGAACACAGCCATCATGACGTACAGACTCTT 564  
 Qy 645 CCCCAAGCTTGGGCTCTTTCCACGAGGCTTCTCTGACGGGCGGAGAGCTTCTCCAG 704  
 Db 555 CCCCAAGCTTGGGCTCTTTCCACGAGGCTTCTCTGACGGGCGGAGAGCTTCTCCAG 624  
 Qy 705 TTGTGGGTCT-GGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGT 763  
 Db 625 TTGTGGGTCTGGGGGTGGTGGCATCTCCCTAAAGAGTGGTCCCTCTGTGGCAAGT 684  
 Qy 764 ATGAAGTCTCAGCTTGTGCTTCTCTCCACAGCAAGAAAGACTGGTCC-TTTTGTGTG 822  
 Db 685 ATGAAGTCTCAGCTTGTGCTTCTCTCCACAGCAAGAAAGACTGGTCTTTTGTGTG 744  
 Qy 823 TTGGAACCCCAAGATC-AGAGGATTTGGAGCCCTGTAAGAAAGAAATGAGAGGATG 880  
 Db 745 TTGGAACCCCAAGATCAGAGGATTTGGAGCCCTGTAAGAAAGAAATGAGAGGATG 804  
 Qy 881 GGGACCTT-GA-CCTGAACGGGCAAGTGTGTG--TGACACAACGGCGTGAAGATGCCA 935  
 Db 805 GGGACCTTGAACCTTGAACGGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 864  
 Qy 936 AACCG 940  
 Db 865 AACCG 869

RESULT 15  
 BQ924736 938 bp mRNA linear EST 20-AUG-2002  
 LOCUS BQ924736  
 DEFINITION AGENCOURT 8853777 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6574110  
 5' mRNA sequence.

ACCESSION BQ924736  
 VERSION BQ924736  
 KEYWORDS EST.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)

REFERENCE  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapds-remail.nih.gov  
 Tissue Procurement: DCTD/DTP/Gazdar  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM2552 row: 1 column: 23  
 High quality sequence step: 629.

## FEATURES

source  
 1..938  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6574110"  
 /tissue\_type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_18"  
 /note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGGCAGAG(G) Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN

Query Match 26.7%; Score 714; DB 5; Length 938;  
 Best Local Similarity 89.7%; Pred. No. 28-173;  
 Matches 874; Conservative 0; Mismatches 47; Indels 53; Gaps 8;

Qy 48 CAGCCGTTCCGGGTTCCGGGCGGGGCGGGAGTGAATCCGATGAGCCGCCAGAGA 107  
 Db 1 CAGCCGTTCCGGGTTCCGGGCGGGGCGGGAGTGAATCCGATGAGCCGCCAGAGA 60  
 Qy 108 AGGCAAGCATGTGCGCGCGCCGCGGAGCTCTGGGAGAGGCTCTGGGCGCGGAGAG 167  
 Db 61 AGGCAAGCATGTGCGCGCGCCGCGGAGCTCTGGGAGAGGCTCTGGGCGCGGAGAG 120  
 Qy 168 GGGCGAGCGGCAAGTCTCTCTGAGGAGGAGGAGTGAACATCGGGCGGAGAGAGGTTG 227  
 Db 121 GGGCGAGCGGCAAGTCTCTCTGAGGAGGAGGAGTGAACATCGGGCGGAGAGAGGTTG 180  
 Qy 228 CGACCTTCTCTTCCCGAGCAATAACTGTCTCTGAGAGTCACTGAAATTTAGTGA 287  
 Db 181 CGACCTTCTCTTCCCGAGCAATAACTGTCTCTGAGAGTCACTGAAATTTAGTGA 240  
 Qy 288 TGAATAATCAGGTCAAGTGAACCTGGAAGATACAGCACAGTGAACAGTATTAAACA 347  
 Db 241 TGAATAATCAGGTCAAGTGAACCTGGAAGATACAGCACAGTGAACAGTATTAAACA 300  
 Qy 348 GCTGAAGGTGTTAAGAGCAGACATGCCCTTACAGACTGGGGATGTGATCTACTGTGT 407  
 Db 301 GCTGAAGGTGTTAAGAGCAGACATGCCCTTACAGACTGGGGATGTGATCTACTGTGT 360  
 Qy 408 GTACAGGAAGATGAACCGGAACACACAGTGCATACCTTATGATCTTTAAGTGA 467  
 Db 361 GTACAGGAAGATGAACCGGAACACACAGTGCATACCTTATGATCTTTAAGTGA 420  
 Qy 468 GCAAGGATGACACAGAATCTTTGAAGCTAACAGGAATGTGTTCCATGGAGACAA 527  
 Db 421 GCAAGGATGACACAGAATCTTT----- 445  
 Qy 528 AGATACCTCAGGTGACAGGTGACGGGCGAGGGCGCATCCCGGGTCCCTCGTGGTGGC 587  
 Db 446 -GATACCTCAGGTGACAGGTGACGGGCGAGGGCGCATCCCGGGTCCCTCGTGGTGGC 504  
 Qy 588 CCCCACTCAGGTGCTTCTTGAAGAAACAAGCCATCAACATGACGTGACGCTCTTCCC 647  
 Db 505 CGCACTCAGGTGCTTCTTGAAGAAACAAGCCATCAACATGACGTGACGCTCTTCCC 564  
 Qy 648 CACAGCTGCGCTCTTCAAGAGCTTCTCTGAGGGGAGAGCGTTCCTCAAGTTG 707  
 Db 565 CACAGCTGCGCTCTTCAAGAGCTTCTCTGAGGGGAGAGCGTTCCTCAAGTTG 624  
 Qy 708 TGGGTCTGGGGGTGGTGGCATCTCCCTTAAAGAGGTGCTCTGTGGGAGAGTGA 766  
 Db 625 TGGGTCTGGGGGTGGTGGCATCTCCCTTAAAGAGGTGCTCTGTGGGAGAGTGA 684  
 Qy 767 AAGTCTCAGCTTCTCTCCTCAAGCTCTCCAGACAGAAAGACTGCTCTTCTGTCGTTGG 826  
 Db 685 AAGTCTCAGCTTCTCTCCTCAAGCTCTCCAGACAGAAAGACTGCTCTTCTGTCGTTGG 744  
 Qy 827 AA-CCCAAGATCAGAGAGATTGAGAGCC-----GTGAAGAAATGAGAGAGATG 880  
 Db 745 AACCCCAAGATCAGAGAGATTGAGAGCCGAGAGAAAGAAATGANNAGAGATGGG 804  
 Qy 881 GGGACCTTGAACCTGAA-----CGGCAAGTGTGTGTGCGACAAACGGCGTAGAAATGGCCA 935  
 Db 805 GGGACCTTGAACCTGAAACGGGCAAGTGTGTGTGTGCGACAAACGGCGTAGAAATGGCCA 864  
 Qy 936 AACCTTCAAG 990  
 Db 865 AACCTTCAAG 924  
 Qy 991 ACGTGAATGAT 1004  
 Db 925 ACGTGAATGAT 938

Sun Jan 16 16:00:06 2005

us-10-048-046-1.rst

Page 16

Search completed: January 15, 2005, 06:17:46  
Job time : 8064 secs

---



Best Local Similarity 3.4%; Pred. No. 0.0081;  
Matches 13; Conservative 219; Mismatches 155; Indels 0; Gaps 0;

```

QY 191 GGAAGCGGAGTGGACCATCGGGCGGAGCAGAGTGGCCACTTCTCCCGACAATA 250
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1438 GTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1379
QY 251 AACTGCTCTGTGACATCACTGATAATGTAGTGAATAAATCATGACGTGACAC 310
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1378 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1319
QY 311 TGGAAATATACGACCACTGGAACAGTATTAACAAGCTGAAGTTGTAAGAACAGA 370
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1318 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1259
QY 371 CATGCCCTTACGACTGTGGATGTATCTAATTGTTGTAAGAAATGAACGAGAC 430
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1258 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1199
QY 431 ACAAGTGCATACCTCTATGATCTTTAAGTAAAAAGAGCATGACACAAGATCT 490
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1198 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1139
QY 491 TTGAAGCTAAACAAGAAATGTCTTCATGGAGCAACAAATACCTCAGGTGACG 550
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1138 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1079
QY 551 GCGAGGCGCGATCCCGGCTCCTC 577
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1078 RRRRRRRRRRRATGCAAGCTCCCTC 1052

```

## RESULT 2

US-09-621-976-16656  
Sequence 16656, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jober, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET. 054PR2  
CURRENT APPLICATION NUMBER: US/09/621, 976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 16656  
LENGTH: 430  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-16656

Query Match 1.7%; Score 46.2; DB 4; Length 430;  
Best Local Similarity 17.4%; Pred. No. 0.039;  
Matches 57; Conservative 126; Mismatches 144; Indels 0; Gaps 0;

```

QY 1528 GGAAGCACGCGCTGTGACCTGCTGCTTCAGCCCATGCCGACGAGCGAGCGC 1587
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 SGRAACCTYYKKKGGSCSRAMCCCYTKYKSCSSYKSGYTKRAMMMKRSCYTSRR 115
QY 1588 GAGCAGGACCGCGGTGTGCGCTCAGCAGGTGTGGGTGTGCTCAGCCTTTCGCA 1647
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 116 MYTTRSTYMMWSMKSGSCSSGSGSCYKKGKKGKSGCMESYMWCCYYRARRMMWK 175
QY 1648 CTGTACTGGGCTGACCCGACCGGCTGCTACGCGCTGCTGCGCCGCTTTGTAGCTC 1707
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 176 GSGCMYTRMMRBMCCCMRRRSMRMRMCMKMSYTYCYSSSMCMARMMKARAKK 235
QY 1708 AACCTGGGTGACAGTGTGTGACGCGCTGTGAACAACACAGCTGACAGTCAATC 1767
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 236 RMCCTYKGGGMMRYCCMKRKGRACTGTTCACTGCCGTGCTGCAAGAGCCCTTC 295
QY 1768 CTGAAGATTACTGCAACGAGGTTTGAATGAAAAACAATGTGACGAGAGCCTC 1827

```

```

DB 236 ACTTCCAGCCCATGTYKAMCSGCMWKRARKKYYMMAMSRMSKYARRRSGWCCYY 355
QY 1828 GTGCTCTCCAGCGGGAGTGTCTG 1854
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 356 KSMMSKMYKSGGAGGCTTCAATG 382

```

## RESULT 3

US-09-807-258-13  
Sequence 13, Application US/09807258  
Patent No. 6670166  
GENERAL INFORMATION:  
APPLICANT: E. I. du Pont de Nemours and Company  
TITLE OF INVENTION: Arthropod Protein Disulfide Isomerases  
FILE REFERENCE: BB-1253 PCT  
CURRENT APPLICATION NUMBER: US/09/807, 258  
CURRENT FILING DATE: 2001-06-11  
PRIOR APPLICATION NUMBER: 60/104, 376  
PRIOR FILING DATE: 1998-10-15  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 13  
LENGTH: 1759  
TYPE: DNA  
ORGANISM: Scolopendra canidens DS  
US-09-807-258-13

Query Match 1.7%; Score 45.8; DB 4; Length 1759;  
Best Local Similarity 49.4%; Pred. No. 0.11; Indels 122; Gaps 0;  
Matches 119; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

```

QY 2439 AGAAMAAGTTCAAGGAGGAGGCAAGTTATCAAAAACATGTTTCAGAGAGGAG 2498
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1514 AGAAAAAGTTGAATGTAAGTGAATTTAGATTTAAATATTTGTTCAATAAAGC 1573
QY 2499 CATAGTTTACGCTACAGAGCTACAAATATCTGCTGCGAAAAACACAGCAT 2558
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1574 ACAATTTTATTTTAAAGGAAATMAAATGTAATTAATTTAATTTAATTT 1633
QY 2559 TTATCTATTTTATTTTAAATAGTTGGTCTATCTCAATATGATTTAATATGAC 2618
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1634 AGAGTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1693
QY 2619 AACTGTAGACAAATATATATTTATTTAATTTAATTTAATTTAATTTAATTT 2678
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1694 AATTAATCGAAATTAATTTTATTTAATTTAATTTAATTTAATTTAATTTA 1753
QY 2679 A 2679
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1754 A 1754

```

## RESULT 4

US-08-924-747-25  
Sequence 25, Application US/08924747  
Patent No. 6063570  
GENERAL INFORMATION:  
APPLICANT: MCGONIGLE, BRIAN  
APPLICANT: O'KEEFE, DANIEL  
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH  
COMPUTER: IBM PC COMPATIBLE



OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95  
SOFTWARE: MICROSOFT WORD VERSION 7.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/924,747  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMETHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CL-1108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 991 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
TISSUE TYPE: SOYBEAN  
IMMEDIATE SOURCE:  
CLONE: SSM.PK0067.G5  
US-08-924-747-25

Query Match 1.7%; Score 45.4; DB 3; Length 991;  
Best Local Similarity 51.2%; Pred. No. 0.099;  
Matches 106; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 2473 AAAAATTGTTTCAGAGAGGAGCATTAAGTTTACAGCCTACAGAGCTACAAATAT 2532  
DB 766 AGAACATTTTAGAGTAGAGTTCATTAATAGTAGATTTTGTGGAAACAATTA 825  
QY 2533 CCTGCTGCTGGGAAACACAGCATTTATCTATTTTATTTAATAGTTTGTGCTT 2592  
DB 826 TCTTGTGTGAGCAAGAGATTTGTCGTTTAAATTAATTAATGCTGATTTGTTGGGT 885  
QY 2593 ATCTTCTAATTAAGATTTAAATGTCACAACTGTAGACAAATTAATTAATTAATTA 2652  
DB 886 ATGGCTATTTTAATTTTAATTAATAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAA 945  
QY 2653 CAATTCACAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAA 2679  
DB 946 AAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAA 972

RESULT 5  
US-09-247-373B-25  
Sequence 25, Application US/09247373B  
Patent No. 6168954  
GENERAL INFORMATION:  
APPLICANT: MCGONIGLE, BRIAN  
APPLICANT: O'KEEFE, DANIEL  
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES  
FILE REFERENCE: CL-1108-A  
CURRENT APPLICATION NUMBER: US/09/247,373B  
CURRENT FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 08/924,747  
PRIOR FILING DATE: 1997-09-05  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 25  
LENGTH: 991  
TYPE: DNA  
ORGANISM: SOYBEAN  
US-09-247-373B-25

Query Match 1.7%; Score 45.4; DB 3; Length 991;  
Best Local Similarity 51.2%; Pred. No. 0.099;  
Matches 106; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 2473 AAAAATTGTTTCAGAGAGGAGCATTAAGTTTACAGCCTACAGAGCTACAAATAT 2532  
DB 766 AGAACATTTTAGAGTAGAGTTCATTAATAGTAGATTTTGTGGAAACAATTA 825  
QY 2533 CCTGCTGCTGGGAAACACAGCATTTATCTATTTTATTTAATAGTTTGTGCTT 2592  
DB 826 TCTTGTGTGAGCAAGAGATTTGTCGTTTAAATTAATTAATGCTGATTTGTTGGGT 885  
QY 2593 ATCTTCTAATTAAGATTTAAATGTCACAACTGTAGACAAATTAATTAATTAATTA 2652  
DB 886 ATGGCTATTTTAATTTTAATTAATAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAA 945  
QY 2653 CAATTCACAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAA 2679  
DB 946 AAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAAAGTGTTCAGTTTAAAAA 972

RESULT 6  
US-09-296-715-25  
Sequence 25, Application US/09296715  
Patent No. 6171839  
GENERAL INFORMATION:  
APPLICANT: MCGONIGLE, BRIAN  
APPLICANT: O'KEEFE, DANIEL  
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE  
TITLE OF INVENTION: ENZYMES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: E.I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 19898  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.50 INCH  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95  
SOFTWARE: MICROSOFT WORD VERSION 7.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/296,715  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMETHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CL-1108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 991 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
TISSUE TYPE: SOYBEAN  
IMMEDIATE SOURCE:  
CLONE: SSM.PK0067.G5  
US-09-296-715-25

Query Match 1.7%; Score 45.4; DB 3; Length 991;  
Best Local Similarity 51.2%; Pred. No. 0.099;  
Matches 106; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 2473 AAAAATTGTTTCAGAGAGGAGCATTAAGTTTACAGCCTACAGAGCTACAAATAT 2532  
DB 766 AGAACATTTTAGAGTAGAGTTCATTAATAGTAGATTTTGTGGAAACAATTA 825

```
QY 2533 CCTGCTGCTGGGAAAACACAGATTATCTATTTTATTTAATAGTTTGTGCTT 2592
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Db 826 TCTTGTGTTGAGCAAGATGTTCTGTTTAAATTAATGATGTTGTTGGGT 885
QY 2593 ATCTTATTAAGATTAAATGTCACAACTGTAGACAAATTAATTAATTAATTA 2652
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Query Match 1.7%; Score 45.2; DB 3; Length 4403765;
Best Local Similarity 52.1%; Pred. No. 10;
Matches 101; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 32 GACACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 91
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3793359 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3793418
QY 92 TGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 151
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3793419 GGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3793478
QY 152 GTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 211
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3793479 GRCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3793538
QY 212 GCGCGAGACGAGGT 225
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3793539 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3793552
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

RESULT 8
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
```

```
QY 32 GACACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 91
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Db 3802537 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3802596
QY 92 TGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 151
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3802597 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3802656
QY 152 GTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 211
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3802657 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3802716
QY 212 GCGCGAGACGAGGT 225
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

Db 3802717 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3802730
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2
```

```
RESULT 9
US-10-140-002-321
; Sequence 321, Application US/10140002
; Patent No. 6725730
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gettisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 321
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-002-321

Query Match 1.6%; Score 42.8; DB 4; Length 783;
Best Local Similarity 54.4%; Pred. No. 0.41; 72; Indels 0; Gaps 0;
Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 2522 GTACACATATCCGCTGCTGGGAAAACACAGATTATCTATTTTATTTAATAG 2581
```

Db 625 GAAGCTACATGAGCCGCTGGGAGAGAGGGTGTCTCCAGAGTTACTTAATA 684  
 QY 2582 GTTGGTCTTATCTTATATAGATTTAATGTCAACAATGTAGCAACAATAATATA 2641  
 Db 685 AGCTTGTATAGTGAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 744  
 QY 2642 TTATATATTACAAATGCAAAAAAAAAAAAAAAAAAAAA 2679  
 Db 745 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 782

## RESULT 10

US-08-910-925-2/c  
 ; Sequence 2, Application US/08910925  
 ; Patent No. 6162601  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Shah, Purvi  
 ; TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: PasteSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/910,925  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0365 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-855-0555  
 ; TELEFAX: 650-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2369 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: FIBRNOT01  
 ; CLONE: 53219  
 ; US-08-910-925-2

Query Match 1.6%; Score 42.6; DB 3; Length 2369;  
 Best Local Similarity 58.1%; Pred. No. 0.83; Mismatches 54; Indels 0; Gaps 0;  
 Matches 75; Conservative 0;

QY 1639 TTCTGCACTGTACTGAGGAGTGCACCCGAGCCGCTGCTACGCGTGCCTGCGCCGCTTT 1698  
 Db 1873 TGTGTCACCTGTACTGAGAGTGTCTACTGGAATCTGAGACTGCTACTACTCTCACTGC 1814  
 QY 1699 TGTGACTCAACTGTGCTGACAGTGTCTGACGCGCTGCTGAACAACAAGCTTACGAG 1758  
 Db 1813 TGAATCTACTTCCACTGCTGCTACTGTTGAAGTCTACTGAAGTCTACTGCTACCTGC 1754  
 QY 1759 TCAGACATC 1767  
 | | | | |

Db 1753 TTGACTTC 1745

## RESULT 11

US-09-252-991A-2236/c  
 ; Sequence 2236, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 2236  
 ; LENGTH: 966  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-2236

Query Match 1.6%; Score 42.4; DB 4; Length 966;  
 Best Local Similarity 54.5%; Pred. No. 0.58; Mismatches 71; Indels 0; Gaps 0;  
 Matches 85; Conservative 0;

QY 74 CGGGATGTGATATCCCATGAGCCGCCGAGAGGAGCAAGCATGTCGCCGCCGCGCAGC 133  
 Db 578 CGCTGCTGTATCCCATGATCGCTGAGCAAGAGGAGGAGCGCTGGCGGAGG 519  
 QY 134 CCGGGAGCGCTCTGCGTCTGAGCGCGGAGAGGAGGAGCGGACGTCCTCTGAGGA 193  
 Db 518 CCGTGGGGGTCCGAGAGCCGCGGAGAGACTTCCGCCGCGGATCTGAGCGCG 459  
 QY 194 AGCGGAGTGCACCATCGGCGGAGAGAGGTTGCG 229  
 Db 458 AGCATGCGTCTGATCGGTCTATGCGCGGCTTGGCG 423

## RESULT 12

US-09-252-991A-2607  
 ; Sequence 2607, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 2607  
 ; LENGTH: 1734  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-09-252-991A-2607

Query Match 1.6%; Score 42.4; DB 4; Length 1734;  
 Best Local Similarity 54.5%; Pred. No. 0.79; Mismatches 71; Indels 0; Gaps 0;  
 Matches 85; Conservative 0;

QY 74 CGGGATGTGATATCCCATGAGCGGCCGAGAGGAGCAAGCATGTCGCCGCCGCGCAGC 133  
 Db 1212 CGCTGCTGTATCCCATGATCGCTGAGCAAGAGGAGGAGCGCTGGCGGAGG 1271  
 QY 134 CCGGGAGCGCTCTGCGTCTGAGCGCGGAGAGGAGGTTGCG 193











[illegible]

RESULT 3  
US-09-764-864-22

```

: Sequence 2, Application US/09764864
: Patent No. US20020132753A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1223
: CURRENT APPLICATION NUMBER: US/09/764,864
: PRIORITY FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 22
: LENGTH: 1311
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1281)
: OTHER INFORMATION: n equals a,t,g, or c
: US-09-764-864-22

```

Query Match	35.3%	Score 945;	DB 9;	Length 1311;
Best Local Similarity	81.1%	Pred. No. 1.7e-243;		
Matches 1267; Conservative	6;	Mismatches 7;	Indels 282;	Gaps 4

QY	56	TCGCGGGTTCGGGCGCGGGGGGGGGAGTGAATCCCGAATGAGCGGCGCCCGAGGAAGGCAAG	115
Db	5	TCGCGGTTTCGGGCGCGGGGCGGGAGATGAAATCCCGAATGAGCGGCGCCCGAGGAAGGCAAGC	64
QY	116	AGTCGCGCGCGCGCGCAGCCTCTGGGGAGACGGCTCTTGCGTCTGGCGCGGAGAGAGGGCGAGC	175
Db	65	AGTCGCGCGCGCGCGCAGCCTCTGGGGAGACGGCTCTTGCGTCTGGCGCGGAGAGAGGGCGAGC	124
QY	176	CGCAGCTCTCTTGAGGAAGCGGGAGTGAACATCGGGCGGAGACGAGGTTGCGACTTT	235
Db	125	CGCAGCTCTCTCTGAGGAAGCGGGAGTGAACATCGGGCGGAGACGAGGTTGCGACTTT	184
QY	236	CGTTCGCCGAGCAATTAACGTGCTCTGAGGATCACTGTGAATTTGTAGTGGATGAAAAT	295
Db	185	CGTTCGCCGAGCAATTAACGTGCTCTGAGGATCACTGTGAATTTGTAGTGGATGAAAAT	244
QY	296	CAGCTCAGGTGACACTGGAAGATCACGAGCCAGTGGAAACGTGATTAACAGCTGAAG	355

Db	245	CAGGTGAGTGAACCTGGAAAGATCCAGACCAAGTGAACAGTATTACARCTGAARG	304
Qy	356	TTGTTAAGAGCAGACATATGCCCTTTAAGACTGGGAGTGTATCTACTTGTGTACGA	415
Db	305	TTGTTAAGAAACAGACATATGCCCTTTAAGACTGGGAGTGTATCTACTTGTGTACAGA	364
Qy	416	AGATGAACCGGAAACACACAGTGGCATCTCTATGAATCTTTAAGTAAAGCAAGCA	475
Db	365	AGATGAACCGGAAACACA-----	382
Qy	476	TGACCAAGATTCCTTTGAAGCTTAAACAAGAAATGTTCCTATGGGACCAAGATACCT	535
Db	383	-----	382
Qy	536	CAGGTGACAGTGCAGAGGCGAGGGCCGATCCCGGATCCCTCGTGTGCGCCCACTC	595
Db	383	-----	382
Qy	596	AGGTGTCTTTAGAGAACACAGCCATCAATGACGTGACAGCTCTTCCCAAGCCT	655
Db	383	-----	382
Qy	656	CGGCTCTTCCAGGAGCTTCTCTGAGGGCGAGCGTTCTCCAGTTGTGGGTCTG	715
Db	383	-----GgTCTG	388
Qy	716	GGGTGTGTGATCTCCCTTAAAGAAAGTGTCCCTCTGTGGCAAGTATGAATCTTCA	775
Db	389	GGGTGTGTGATCTCCCTTAAAGAAAGTGTCCCTCTGTGGCAAGTATGAATCTTCA	448
Qy	776	GCTTTGCTCAGCTCTCTCCAGACAGAAAGACTGCGTCTTTTGTCTGTGGAACCCAGG	835
Db	449	GCTTTGCTCAGCTCTCTCCAGACAGAAAGACTGCGTCTTTTGTCTGTGGAACCCAGG	508
Qy	836	ATCAGAGAGATTTTGAAGCCGTGAAAGAAATAAGAGAGATGGGAGCCTTGAACCTGA	895
Db	509	ATCAGAGAGATTTTGAAGCCGTGAAAGAAATAAGAGAGATGGGAGCCTTGAACCTGA	568
Qy	896	ACGGGCAAGTTTGTGGTGCACAACCGCGTAGAAATGCCAAACCGTCCACGAGACGTCA	955
Db	569	ACGGGCAAGTTTGTGGTGCACAACCGCGTAGAAATGCCAAACCGTCCACGAGACGTCA	628
Qy	956	GAGCAGCGGCTGGGAAGCCAGACAAGATGAGAGAGACGCTGACATGCAATCTTCCAGG	1015
Db	629	GAGCAGCGGCTGGGAAGCCAGACAAGATGAGAGAGACGCTGACATGCAATCTTCCAGG	688
Qy	1016	ACCTGTCTGACAGACTGCGTAGTTTGTGCAGCCCTTGCAATGACACGTTCTGTGCGGCTTGCT	1075
Db	689	ACCTGTCTGACAGACTGCGTAGTTTGTGCAGCCCTTGCAATGACACGTTCTGTGCGGCTTGCT	748
Qy	1076	ACTCGGGGTGGATGGAGGCGTCTGTCCTGATCTTCACTGCGCGGTGCCGTGGAGCGGA	1135
Db	749	ACTCGGGGTGGATGGAGGCGTCTGTCCTGATCTTCACTGCGCGGTGCCGTGGAGCGGA	808
Qy	1136	TCTGTAAAAACACATCTCTCAACAACCTCGTGAAGCAATACCTATCCAGATCCAGACA	1195
Db	809	TCTGTAAAAACACATCTCTCAACAACCTCGTGAAGCAATACCTATCCAGATCCAGACA	868
Qy	1196	AGAGTCCGAGTGAAGAAAGATGTGCAAAGTATGATGCCAAGAAATAAATCACTCAAGCA	1255
Db	869	AGAGTCCGAGTGAAGAAAGATGTGCAAAGTATGATGCCAAGAAATAAATCACTCAAGCA	928
Qy	1256	TGCTGCAACCCAAAGTCAGCGGCTTTTCTGTATGAAGAAAGGAGTTCAAGAGACCTGC	1315
Db	929	TGCTGCAACCCAAAGTCAGCGGCTTTTCTGTATGAAGAAAGGAGTTCAAGAGACCTGC	988
Qy	1316	TGAGAGTGTCAAGCGTTGACAGTAGATCTTCAGACATTAGCCAGCCATATCGTGTGTGCC	1375
Db	989	TGAGAGTGTCAAGCGTTGACAGTAGATCTTCAGACATTAGCCAGCCATATCGTGTGTGCC	1048
Qy	1376	GGCAGTGTCTGAGTACAGAAAGCAAGCGGCGCAGCTCCCACTGCCAGACCCGAGG	1435

Db	1049	GGCAGTGTCTTGATGACAGAAAGGACAGGCGCGCAGCTTCCCATCTGCCACGAGACCCGAGG	1108
Qy	1436	GGGAGCCGAGAGACCCCAACAAGGCTCTGGGGGATATGCACCTTCCACGTCTCGTCACTGTACAG	1495
Db	1109	GCGAGCCGAGAGGCCCCACAGAGCTCTGGGGGAGATGCACCTTCACTGTCTCGTCACTGTACAG	1168
Qy	1496	CAGCAGTCAGGATTAAGTGTGCTCCCTTGCAGAGGAAGCCAGGCTGTGCACCTGTGCT	1555
Db	1169	CAGT-----CAGGATTAAGTGTGCTCTTGTGCAGAGGAAGCCAGGCTGTGCACCTGTGCT	1224
Qy	1556	TCGAGCCCAATGCCGACCGGAGAGCGAGCGGAGCGGACAGAG--CCCGCTGTGCGCCCTTCAG	1614
Db	1225	T-CAGCCCATGCGCCGACCGGAGATGTGGAACCGGAACAGAGCCCGCTGTGCGSCCTTAA	1283
Qy	1615	CA 1616	
Db	1284	CA 1285	

RESULT 4  
US-10-09

```

?       ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?      ?
? Sequence 160, Application US/10094749
? Publication No. US20030219741A1
? GENERAL INFORMATION:
? APPLICANT: ISOGAI, TAKAO
? APPLICANT: SUGIYAMA, TOMOYASU
? APPLICANT: OTSUKI, TETSUJI
? APPLICANT: WAKAMATSU, AI
? APPLICANT: SATO, HIROYUKI
? APPLICANT: ISHII, SHIZUKO
? APPLICANT: YAMAMOTO, JUN-ICHI
? APPLICANT: ISONO, YUUKO
? APPLICANT: HIO, YURI
? APPLICANT: OTSUKA, KAORU
? APPLICANT: NAGAI, KEIICHI
? APPLICANT: IRIE, RYOTARO
? APPLICANT: TAMECHIKA, ICHIRO
? APPLICANT: SEKI, NAOHICO
? APPLICANT: YOSHIKAWA, TSUTOMU
? APPLICANT: OTSUKA, MOTOUKI
? APPLICANT: NAGAHARI, KENJI
? APPLICANT: MASUHO, YASUHIKO
? TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
? FILE REFERENCE: 08435/0160
? CURRENT APPLICATION NUMBER: US/10/094,749
? PRIOR FILING DATE: 2002-03-12
? PRIOR APPLICATION NUMBER: 60/350,435
? PRIOR FILING DATE: 2002-01-24
? PRIOR APPLICATION NUMBER: JP 2001-328381
? PRIOR FILING DATE: 2001-09-14
? NUMBER OF SEQ ID NOS: 3381
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 160
? LENGTH: 2186
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-094-749-160

```

	Query Match	32.7%	Score 876.21	DB 15	Length 2186
	Similarity	99.4%	Pred. No.7.8e-25		
	Matches	890	Conservative	0	Mismatches 3
					Indels 2
					Gaps 1
QY	1772	AGAAATTACTGCAACGACAGGTTTGAATGAAAACATGTTGACGAGAGCCTCGGG	1831		
Db	671	AGAAATTACTGCAACGACAGGTTGACATGAAAAACATGTTGACGAGAGCCTCGGG	730		
QY	1832	CTCTCCACCGGGAGCTGTTTCTGCTGTCTGATTCAGAGTCACCGGAGACACCGTTCTGT	1891		
Db	731	CTCTCCACCGGGAGCTGTTTCTGCTGTCTGATTCAGAGTCACCGGAGACACCGTTCTGT	790		
QY	1892	GTTTCTGCTGAGGCGCTCGGACGCTCCGCTGAGCTGACCTATTCAGATTCGGAGAACTTC	1951		
Db	791	GTTTCTGCTGAGGCGCTCGGACGCTCCGCTGAGCTGACCTATTCAGATTCGGAGAACTTC	850		

QY	1952	CTGCTTCGCAATGCGCATGCGCGCTAAATCCGCGCTGACCTGCTACCTGGGCGCGTACT	2011
Db	851	CTGCTTCGCAATGCGCATGCGCGCTAAATCCGCGCTGACCTGCTACCTGGGCGCGTACT	910
QY	2012	GCCGCACCTCAGAGTAAAGCTCACCACGCGCATGAATTCATCATATCTGTGAACAGACAA	2071
Db	911	GCCTGACCTCAGAGTAAAGCTCACCACGCGCATGAATTCATCATATCTGTGAACAGACAA	970
QY	2072	GCTTCAAAAATAAGCATCCAGAGCGCCCTGAGCACTTTACGACTGAGGTGAAGAG	2131
Db	971	GCTTCAAAAATAAGCATCCAGAGCGCCCTGAGCACTTTACGACTGAGGTGAAGAG	1030
QY	2132	CGTGTATTTTAAATATACAGAGCAAGCAAGTCATAGTGTATTTTACAGCCCTCTGAGGGAAG	2191
Db	1031	CGTGTATTTTAAATATACAGAGCAAGCAAGTCATAGTGTATTTTACAGCCCTCTGAGGGAAG	1090
QY	2192	GGAAGCAGAGGCTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGACTTTTACCTCT	2251
Db	1091	GGAAGCAGAGGCTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGACTTTTACCTCT	1148
QY	2252	GAGTGAAGCCCTCCGAGAGCCCGGGGGCCGAGCCCGCCTCTCTGTGTGAAGCGCTGGGC	2311
Db	1149	GAGTGAAGCCCTCCGAGAGCCCGGGGGCCGAGCCCGCCTCTCTGTGTGAAGCGCTGGGC	1208
QY	2312	AGGCGTCGTGTGGCATCAGCAGACAGACGAAAGCTTTCTGTAACTTGGCGCGCTCCG	2371
Db	1209	AGGCGTCGTGTGGCATCAGCAGACAGACGAAAGCTTTCTGTAACTTGGCGCGCTCCG	1268
QY	2372	CGAGAGGGGGAGTTTGCTCTTTGTATCAATTTCCGAAACTACAGTTAAAGCAGAAATC	2431
Db	1269	CGAGAGGGGGAGTTTGCTCTTTGTATCAATTTCCGAAACTACAGTTAAAGCAGAAATC	1328
QY	2432	TGTTTTACGAAAAAGTTTCAAGGGAGAAAGGCAAGTTTATCAAAAAATTGTTTACAGAG	2491
Db	1329	TGTTTTACGAAAAAGTTTCAAGGGAGAAAGGCAAGTTTATCAAAAAATTGTTTACAGAG	1388
QY	2492	AAGGAGACATTAAGTTTACAGGCTTACAGAGCGTACACATATCTCGTGTGGGAAAAACA	2551
Db	1389	AAGGAGACATTAAGTTTACAGGCTTACAGAGCGTACACATATCTCGTGTGGGAAAAACA	1448
QY	2552	CAGCAATTTATCTAATTTTATTTAATAGTTTGGTGCTATCTCTGTAAATGAATTTAA	2611
Db	1449	CAGCAATTTATCTAATTTTATTTAATAGTTTGGTGCTATCTCTGTAAATGAATTTAA	1508
QY	2612	ATGTCACAAACGTGAGCAAAATATATTAATTTAATTTAACAATTGACAACAAA	2666
Db	1509	ATGTCACAAACGTGAGCAAAATATATTAATTTTAAATTTTACAAATGACTAATAA	1563

RESULT 5  
116-00-76

```

US-02-764-664-463
; Sequence 483, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT FILING DATE: 2001-01-17
; PRIORITY APPLICATION NUMBER: US/09/764,864
; PRIORITY APPLICATION data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 483
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (13)
;

```

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (623)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (640)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (670)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (687)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (690)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-483

```

Query Match 20.0%; Score 536.6; DB 9; Length 693;

Best Local Similarity 92.5%; Pred.No. 1.3e-133; Matches 608; Conservative 7; Mismatches 13; Indels 29; Gaps 4;

```

QY 1044 GCCCTGATGACAGAGTTCTGCGCGCTTGTACTGCGGCTGATGAGCGCTCGTCCCT 1103
DB 14 GCCCTGATGACAGAGTTCTGCGCGCTTGTACTGCGGCTGATGAGCGCTCGTCCCT 73
QY 1104 GTGTCTTACCTGCGCTGTCTGCGGAGGATCTGTAAACCAATCTTCAACAACCT 1163
DB 74 GTGTCTTACCTGCGCTGTCTGCGGAGGATCTGTAAACCAATCTTCAACAACCT 133
QY 1164 CGTGAAGATCATCTATCCATGACATCCAGAAAGTGGAGTGAAGAAAGTGTCAAG 1223
DB 134 CGTGAAGATCATCTATCCATGACATCCAGAAAGTGGAGTGAAGAAAGTGTCAAG 193
QY 1224 TATGATGACAGAAATTAATCACTCAAGACATGCTGACGCCAAAGTCAAGCGCTCTT 1283
DB 194 TATGATGACAGAAATTAATCACTCAAGACATGCTGACGCCAAAGTCAAGCGCTCTT 253
QY 1284 TTCTGATGAAGAAAGGAGTTCAAGAGACCTGCTGAGCTGTAGACGTTGACATGATC 1343
DB 254 TTCTGATGAAGAAAGGAGTTCAAGAGACCTGCTGAGCTGTAGACGTTGACATGATC 313
QY 1344 CTGACATCTAGACAGCCATCATGCTGAGCGGAGTGTCTGATGACAGAGGACAGC 1403
DB 314 CTGACATCTAGACAGCCATCATGCTGAGCGGAGTGTCTGATGACAGAGGACAGC 373
QY 1404 GCGGAGAGCTCCCACTGCGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1463
DB 374 GCGGAGAGCTCCCACTGCGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 433
QY 1464 GATGACAGCTTCCAGCTGCTGAGCTGACAGACAGAGTTACGTTGCTGCTCT 1523
DB 434 GATGACAGCTTCCAGCTGCTGAGCTGACAGACAGAGTTACGTTGCTGCTCT 489
QY 1524 GCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1583
DB 490 GCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548
QY 1584 GCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1643
DB 549 -CGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG -TAMCMTTSTG 584
QY 1644 CCACTGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1700
DB 585 CCACTGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 641

```

## RESULT 6

```

US-09-918-995-2180
; Sequence 2180, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysreg, Inc.

```

```

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: FROM VARIOUS CDNA LIBRARIES
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2180
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (1) .. (476)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2180

```

Query Match 14.7%; Score 394.4; DB 10; Length 476;

Best Local Similarity 96.7%; Pred.No. 2.2e-95; Matches 413; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

```

QY 2046 ATTCAATCATATCTGTGAACAGACAGAGTTCAAAATCAAGCATCCAGAGGCGCTGACA 2105
DB 52 ATTCAATCATATCTGTGAACAGACAGAGTTCAAAATCAAGCATCCAGAGGCGCTGACA 111
QY 2106 GCTTTCAGACTGAGAGGTGAAGAGAGCGTGTTTTAAATACAGACAAGCACAGTCAAG 2165
DB 112 GCTTTCAGACTGAGAGGAAGAGAGCGTGTTTTAAATACAGACAAGCACAGTCAAG 171
QY 2166 GTGTTTTCAGAGGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2225
DB 172 GTGTTTTCAGAGGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 231
QY 2226 CTCTTCTGTGAGAGGCTTTTAACTCTGAGTGAAGACCTCCCAAGAGGCGGAGGAGG 2285
DB 232 CTCTTCTGTGAGAGGCTTTTAACTCTGAGTGAAGACCTCCCAAGAGGCGGAGGAGG 289
QY 2286 GCCCGGCTCTCTGTGAGAGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2345
DB 290 GACCGGCTCTCTGTGAGAGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 349
QY 2346 CTTTCTGTGAACAGGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2405
DB 350 CTTTCTGTGAACAGGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 409
QY 2406 CCGAACTACAGTTAAAGAGAGTGTGTTTTCAGAAAGATTCAAGGAGGAGGAGG 2465
DB 410 CCGAACTACAGTTAAAGAGAGTGTGTTTTCAGAAAGATTCAAGGAGGAGGAGG 469
QY 2466 GTTTATC 2472
DB 470 GTTTATC 476

```

## RESULT 7

```

US-09-764-864-322
; Sequence 322, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentm Ver. 2.0
; SEQ ID NO 322
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens

```



Query Match	2.2%;	Score 60;	DB 10;	Length 60;
Best Local Similarity	100.0%;	Pred. No.	8.6e-06;	

```

? GENERAL INFORMATION:
? APPLICANT: La Rosa Thomas J
? APPLICANT: Kovallik David K
? APPLICANT: Zhou Yihua
? APPLICANT: Cao Yongwei
? TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
? FILE REFERENCE: 38-21 (53223)B
? CURRENT APPLICATION NUMBER: US/10/424,599
? CURRENT FILING DATE: 2003-04-28
? NUMBER OF SEQ ID NOS: 285684
? SEQ ID NO 95539
? LENGTH: 1255
? TYPE: DNA
? ORGANISM: Glycine max
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (1)..(1255)
? OTHER INFORMATION: unsure at all n locations
? FEATURE:

```





***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: January 15, 2005, 04:02:42 ; Search time 11079 Seconds

(without alignments)  
11435.067 Million cell updates/sec

Title: US-10-048-046-1

Perfect score: 2679

Sequence: 1 agaatctgcgcagcagcgcg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: OLIGO-NTC  
Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 20 *or greater*

Total number of hits satisfying chosen parameters: 72274

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_stb:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2679	100.0	2679	9 AF170724	Homo sapi
2	2133	79.6	2639	6 AX877194	Sequence
3	2133	79.6	2639	6 BD156534	Primer fo
4	2133	79.6	2639	9 AK027687	Homo sapi
5	2031	75.8	3189	9 BC012072	Homo sapi
6	1618	60.4	3181	6 AX405642	Sequence
7	1465	54.7	3138	6 AX877222	Sequence
8	1465	54.7	3138	6 BD156548	Sequence
9	1465	54.7	3138	6 AK001658	Primer fo
10	944	35.2	2448	6 AK834949	Sequence
11	944	35.2	2448	6 AK097671	Homo sapi
12	930	34.7	6235	6 BD183291	Novel gen
13	909	33.9	2257	9 AK090948	Homo sapi
14	633	23.6	2186	6 AX713476	Sequence
15	633	23.6	2186	9 AK054917	Homo sapi
16	483	18.0	1085	6 CQ719982	Sequence
17	478	17.8	816	6 AX868758	Sequence
18	478	17.8	816	6 BD148820	Primer fo
19	468	17.5	824	6 AX868776	Sequence

20	468	17.5	824	6 BD148838	Primer fo
21	426	15.9	518	6 AX873789	Sequence
22	426	15.9	518	6 BD153851	Primer fo
23	386	14.4	1810	9 AC127070	Homo sapi
24	386	14.4	181438	2 AC023047	Homo sapi
25	218	8.1	181438	2 AC023047	Homo sapi
26	98	3.7	756	11 BV180578	sgm11034
27	97	3.6	601	11 BV168191	sgm7638
28	78	2.9	617	6 AX385952	Sequence
29	60	2.2	60	6 CQ356585	Sequence
30	38	1.4	158897	2 AC118260	Mus muscu
31	37	1.4	3194	10 BC049792	Mus muscu
32	37	1.4	321708	2 AC087142	Mus muscu
33	37	1.4	338116	2 AC087159	Mus muscu
34	34	1.3	202792	2 AC125898	Mus muscu
35	32	1.2	601	11 BV002855	Mus muscu
36	29	1.1	124973	8 AC124951	Mus muscu
37	28	1.0	146516	2 AC079981	Mus muscu
38	28	1.0	218226	2 AC126634	Rattus no
39	28	1.0	228987	2 AC111838	Rattus no
40	28	1.0	274796	2 AC097665	Rattus no
41	27	1.0	167166	2 AC117568	Mus muscu
42	27	1.0	167938	10 AC115904	Mus muscu
43	27	1.0	221753	10 AC140355	Mus muscu
44	27	1.0	230221	10 AC123851	Mus muscu
45	27	1.0	234393	10 AL772311	Mouse DNA
46	27	1.0	239783	2 AC102564	Mus muscu
47	27	1.0	266632	2 AC129762	Rattus no
48	27	1.0	269117	2 AC115400	Rattus no
49	27	1.0	322833	27 AC131863	Rattus no
50	26	1.0	444	8 ATMS29544	Arabidops
51	26	1.0	860	6 B77096	Sequence 18
52	26	1.0	860	8 CACORAT16	X70998 C. avellana
53	26	1.0	3144	3 AY061591	Drosophila
54	26	1.0	61165	10 AL669936	Mouse DNA
55	26	1.0	68474	2 AC025193	Mus muscu
56	26	1.0	94784	9 AL355309	Human DNA
57	26	1.0	102008	2 AC016552	Homo sapi
58	26	1.0	113950	8 AC004687	Arabidops
59	26	1.0	134347	9 AL356295	Human DNA
60	26	1.0	149509	5 BX005028	Zebrafish
61	26	1.0	152968	2 AC073669	Mus muscu
62	26	1.0	155451	2 BX942846	Danio rer
63	26	1.0	157262	10 AC121945	Mus muscu
64	26	1.0	168473	5 BX649544	Zebrafish
65	26	1.0	176196	5 AL928676	Zebrafish
66	26	1.0	179683	2 AC115823	Mus muscu
67	26	1.0	182082	2 AC092842	Homo sapi
68	26	1.0	185951	2 AC024386	Homo sapi
69	26	1.0	198939	2 BX957279	Danio rer
70	26	1.0	201467	9 AC093650	Homo sapi
71	26	1.0	201949	2 AC109165	Mus muscu
72	26	1.0	205547	10 AC073938	Mus muscu
73	26	1.0	222149	2 AC115978	Mus muscu
74	26	1.0	222932	2 AC137869	Mus muscu
75	26	1.0	231676	2 AC117155	Rattus no
76	26	1.0	231895	2 AC127043	Rattus no
77	26	1.0	232876	10 AC140338	Mus muscu
78	26	1.0	236350	10 AC020971	Mus muscu
79	26	1.0	239859	2 AC119482	Rattus no
80	26	1.0	241112	2 AC106261	Rattus no
81	26	1.0	244525	9 AC008489	Homo sapi
82	26	1.0	257666	2 AC121225	Rattus no
83	26	1.0	268208	2 AC106249	Rattus no
84	26	1.0	268524	2 AC106223	Rattus no
85	25	0.9	442	8 ATMS29842	Arabidops
86	25	0.9	468	5 BC053809	Xenopus 1
87	25	0.9	870	8 AP425601	Lycopodium
88	25	0.9	1085	6 AR321631	Sequence
89	25	0.9	1400	3 AY069274	Drosophila
90	25	0.9	1468	10 BC025901	Mus muscu
91	25	0.9	1583	9 AB048887	Macaca fa
92	25	0.9	1963	6 AX406978	Sequence

C 93	25	0.9	2000	6	AX508642	AX508642 Sequence	166	24	0.9	2005	8	AY072735	AY072735 Antirrhin
C 94	25	0.9	2000	6	AX652504	AX652504 Sequence	167	24	0.9	2126	10	BC072584	BC072584 Mus muscu
C 95	25	0.9	2415	5	BC009894	BC009894 Homo sapi	168	24	0.9	2215	9	AF086790	AF086790 Homo sapi
C 96	25	0.9	2433	5	BC053227	BC053227 Dantio rer	169	24	0.9	2243	3	AK112541	AK112541 Ciona int
C 97	25	0.9	2569	8	AP000734	AP000734 Arabidops	170	24	0.9	2280	3	AY052085	AY052085 Drosophil
C 98	25	0.9	61779	2	AC126301	AC126301 Rattus no	171	24	0.9	2335	9	BC040064	BC040064 Homo sapi
C 99	25	0.9	77129	2	AC142108_3	Continuation (4 of	172	24	0.9	2596	3	PPADR	PPADR
C 100	25	0.9	77832	9	AL550039	AL550039 Human DNA	173	24	0.9	2654	9	BC030647	BC030647 Homo sapi
C 101	25	0.9	78378	8	AC007260	AC007260 Arabidops	174	24	0.9	3157	6	AR363205	AR363205 Sequence
C 102	25	0.9	87560	6	AX652128	AX652128 Sequence	175	24	0.9	3241	3	PPADRA	PPADRA
C 103	25	0.9	95769	6	AC013430	AC013430 Genomic s	176	24	0.9	3923	9	HSZNF2	HSZNF2
C 104	25	0.9	95766	8	AC005967	AC005967 Arabidops	177	24	0.9	4745	8	PPAERTHBA	PPAERTHBA
C 105	25	0.9	96046	8	AC005967	AC005967 Arabidops	178	24	0.9	8221	8	AP098632	AP098632 Arabidops
C 106	25	0.9	107502	9	AC112200	AC112200 Homo sapi	179	24	0.9	22307	2	AC015297	AC015297 Drosophil
C 107	25	0.9	110000	2	AC142108_2	Continuation (3 of	180	24	0.9	32885	10	AF198490	AF198490 Homo sapi
C 108	25	0.9	110000	3	AC116305_0	AC116305 Dictyoste	181	24	0.9	33036	9	BX255880	BX255880 Mouse DNA
C 109	25	0.9	110804	9	AC004748	AC004748 Homo sapi	182	24	0.9	46626	2	AC007293	AC007293 Arabidops
C 110	25	0.9	117542	9	AC015727	AC015727 Homo sapi	183	24	0.9	47687	2	AC136435	AC136435 Homo sapi
C 111	25	0.9	137628	10	AL672153	AL672153 Mouse DNA	184	24	0.9	51330	2	AC101066	AC101066 Mus muscu
C 112	25	0.9	138859	9	AL359076	AL359076 Human DNA	185	24	0.9	58143	2	AC137177	AC137177 Rattus no
C 113	25	0.9	140148	9	AL354916	AL354916 Human DNA	186	24	0.9	58780	2	AC005085	AC005085 Homo sapi
C 114	25	0.9	140667	2	AC108904	AC108904 Felis cat	187	24	0.9	59605	2	AC103926	AC103926 Mus muscu
C 115	25	0.9	143914	2	AC022711	AC022711 Homo sapi	188	24	0.9	62455	5	CR356237	CR356237 Zebrafish
C 116	25	0.9	144695	2	AC012438	AC012438 Homo sapi	189	24	0.9	66624	2	AC101520	AC101520 Mus muscu
C 117	25	0.9	150485	2	AC012292	AC012292 Homo sapi	190	24	0.9	66860	9	AC129782	AC129782 Homo sapi
C 118	25	0.9	151210	2	AC139591	AC139591 Rattus no	191	24	0.9	72891	8	AB010692	AB010692 Arabidops
C 119	25	0.9	151432	2	AC083966	AC083966 Homo sapi	192	24	0.9	78932	5	BX649412	BX649412 Zebrafish
C 120	25	0.9	154460	10	AC124357	AC124357 Mus muscu	193	24	0.9	84203	8	AC005106	AC005106 Genomic s
C 121	25	0.9	154482	8	AC144431	AC144431 Medicago	194	24	0.9	84294	8	AB019229	AB019229 Arabidops
C 122	25	0.9	156765	2	AC026244	AC026244 Homo sapi	195	24	0.9	85791	8	AB016888	AB016888 Arabidops
C 123	25	0.9	157113	2	AC016337	AC016337 Homo sapi	196	24	0.9	87102	2	AL606983	AL606983 Dantio rer
C 124	25	0.9	161498	9	AC113376	AC113376 Homo sapi	197	24	0.9	90857	2	AC119165	AC119165 Mus muscu
C 125	25	0.9	162394	10	AL683891	AL683891 Mouse DNA	198	24	0.9	93845	8	AC004684	AC004684 Arabidops
C 126	25	0.9	165087	9	AC145772	AC145772 Pan trogl	199	24	0.9	94487	8	AC002291	AC002291 Arabidops
C 127	25	0.9	165087	9	AC145772	AC145772 Pan trogl	200	24	0.9	96442	8	AC002291	AC002291 Arabidops
C 128	25	0.9	171464	9	AC025519	AC025519 Homo sapi	201	24	0.9	97774	5	AL591399	AL591399 Zebrafish
C 129	25	0.9	171413	5	AL929315	AL929315 Zebrafish	202	24	0.9	100806	8	AL591399	AL591399 Zebrafish
C 130	25	0.9	178351	5	AC099542	AC099542 Homo sapi	203	24	0.9	103576	8	YBP8H12	YBP8H12
C 131	25	0.9	178901	5	BX649265	BX649265 Zebrafish	204	24	0.9	103520	8	AC116820	AC116820 Mus muscu
C 132	25	0.9	179115	5	BX333449	BX333449 Zebrafish	205	24	0.9	109432	8	AC136973	AC136973 Medicago
C 133	25	0.9	182405	10	AC124978	AC124978 Mus muscu	206	24	0.9	109922	2	AC01671	AC01671 Homo sapi
C 134	25	0.9	191555	2	CR316778	CR316778 Dantio rer	207	24	0.9	110000	9	AC005079_3	AC005079_3 Continuation (4 of
C 135	25	0.9	191750	2	AC024113	AC024113 Mus muscu	208	24	0.9	118711	9	AC105272	AC105272 Homo sapi
C 136	25	0.9	196755	2	AC140779	AC140779 Mus muscu	209	24	0.9	120364	8	AC140915	AC140915 Medicago
C 137	25	0.9	198793	2	AC117858	AC117858 Rattus no	210	24	0.9	120440	5	BX000467	BX000467 Zebrafish
C 138	25	0.9	203133	2	AC181816	AC181816 Sus scrof	211	24	0.9	129756	9	HS875H10	HS875H10 Human DNA
C 139	25	0.9	206191	10	AC124535	AC124535 Mus muscu	212	24	0.9	133873	8	AC130800	AC130800 Medicago
C 140	25	0.9	209686	2	AC092526	AC092526 Papio anu	213	24	0.9	134001	10	AC096863	AC096863 Genomic s
C 141	25	0.9	213666	2	AC020837	AC020837 Mus muscu	214	24	0.9	136333	8	AC131248	AC131248 Medicago
C 142	25	0.9	213704	10	AL732572	AL732572 Mouse DNA	215	24	0.9	142475	9	AC073525	AC073525 Homo sapi
C 143	25	0.9	219250	9	AC068254	AC068254 Homo sapi	216	24	0.9	142418	2	AC144468	AC144468 Carls tam
C 144	25	0.9	223805	2	AC079475	AC079475 Mus muscu	217	24	0.9	146039	5	AL627132	AL627132 Zebrafish
C 145	25	0.9	224158	2	AC093984	AC093984 Rattus no	218	24	0.9	147986	2	AC102936	AC102936 Mus muscu
C 146	25	0.9	230755	2	AC024618	AC024618 Mus muscu	219	24	0.9	150194	5	BX649510	BX649510 Zebrafish
C 147	25	0.9	231507	2	AC137434	AC137434 Rattus no	220	24	0.9	151277	9	AF268619	AF268619 Homo sapi
C 148	25	0.9	232613	9	AC008679	AC008679 Homo sapi	221	24	0.9	151281	9	AC103680	AC103680 Homo sapi
C 149	25	0.9	239053	2	AC129055	AC129055 Rattus no	222	24	0.9	157000	2	AC026511	AC026511 Homo sapi
C 150	25	0.9	246443	2	AC128415	AC128415 Rattus no	223	24	0.9	157195	2	AC122770	AC122770 Mus muscu
C 151	25	0.9	253774	10	AC102062	AC102062 Mus muscu	224	24	0.9	159324	2	AC115002	AC115002 Mus muscu
C 152	25	0.9	258688	2	AC097024	AC097024 Rattus no	225	24	0.9	159694	2	AC138037	AC138037 Rattus no
C 153	25	0.9	273812	2	AC106441	AC106441 Rattus no	226	24	0.9	164612	2	CR361565	CR361565 Dantio rer
C 154	25	0.9	349418	3	CR382398	CR382398 Plasmodi	227	24	0.9	164639	10	AC124697	AC124697 Homo sapi
C 155	24	0.9	80	3	DR081584	DR081584	228	24	0.9	165405	10	AC119839	AC119839 Mus muscu
C 156	24	0.9	80	3	DR081586	DR081586	229	24	0.9	165405	10	AC119839	AC119839 Mus muscu
C 157	24	0.9	199	6	CQ466844	CQ466844 Sequence	230	24	0.9	165978	10	AC127431	AC127431 Mus muscu
C 158	24	0.9	291	6	CQ488160	CQ488160 Sequence	231	24	0.9	166991	2	AC015952	AC015952 Homo sapi
C 159	24	0.9	320	6	CQ676188	CQ676188 Sequence	232	24	0.9	167441	2	AC092528	AC092528 Mus muscu
C 160	24	0.9	413	6	CQ517947	CQ517947 Sequence	233	24	0.9	168354	2	AC133867	AC133867 Mus muscu
C 161	24	0.9	525	6	CQ529064	CQ529064 Sequence	234	24	0.9	168861	2	CR318590	CR318590 Dantio rer
C 162	24	0.9	844	9	HDMKAP1A	HDMKAP1A	235	24	0.9	169083	2	CR352230	CR352230 Dantio rer
C 163	24	0.9	963	8	AY061024	AY061024 Drosophil	236	24	0.9	169313	2	AC112224	AC112224 Mus muscu
C 164	24	0.9	1585	3	AK115533	AK115533 Ciona int	237	24	0.9	169659	9	AC024933	AC024933 Homo sapi
C 165	24	0.9	2003	3	AK115533	AK115533 Ciona int	238	24	0.9	169659	9	AC024933	AC024933 Homo sapi

239	24	0.9	170687	2	CR589948	Danio rer	c 312	24	0.9	261373	2	AC096416	AC096416 Rattus no
240	24	0.9	170691	9	AC025253	Homo sapi	c 313	24	0.9	272714	2	AC129447	AC129447 Rattus no
241	24	0.9	171994	9	AC147516	Rattus no	c 314	24	0.9	319520	2	BX936378	BX936378 Danio rer
242	24	0.9	172077	10	BX005171	Mouse DNA	c 315	24	0.9	348034	3	CR382400	CR382400 Plasmodu
243	24	0.9	173085	2	CR628344	Danio rer	c 316	24	0.9	349881	6	AX646903	AX646903 Sequence
244	24	0.9	173309	5	BX005313	Human DNA	c 317	24	0.9	160	6	CQ709178	CQ709178 Sequence
245	24	0.9	173719	5	BX005313	Zebrafish	c 318	23	0.9	229	6	ATHS21321	ATHS21321 Arabidops
246	24	0.9	174383	2	AC012053	Homo sapi	c 319	23	0.9	209	8	CO467200	CO467200 Sequence
247	24	0.9	177471	3	AC008361	Drosophila	c 320	23	0.9	234	11	CR378769	CR378769 Arabidops
248	24	0.9	177882	9	AC013402	Homo sapi	c 321	23	0.9	374	6	AX521999	AX521999 Sequence
249	24	0.9	179026	2	CR352332	Danio rer	c 322	23	0.9	443	6	AR426075	AR426075 Sequence
250	24	0.9	179668	10	AC124339	Mouse muscu	c 323	23	0.9	443	6	AX986769	AX986769 Sequence
251	24	0.9	179762	2	AC013519	Homo sapi	c 324	23	0.9	443	6	BD121628	BD121628 EST and e
252	24	0.9	180944	9	AC084855	Homo sapi	c 325	23	0.9	485	11	BY174934	BY174934 sqmm75454
253	24	0.9	182073	5	BX119973	Zebrafish	c 326	23	0.9	530	10	MMU504717	MMU504717 Mus muscu
254	24	0.9	183965	2	AC150567	Bos tauru	c 327	23	0.9	563	11	G96448	G96448 B2086764RD
255	24	0.9	184034	10	AC125163	Mouse muscu	c 328	23	0.9	571	6	AX347755	AX347755 Sequence
256	24	0.9	185316	9	AP002797	Homo sapi	c 329	23	0.9	576	3	AY189223	AY189223 Aedes aeg
257	24	0.9	185562	2	AC109230	Mouse muscu	c 330	23	0.9	582	11	BY179873	BY179873 sqmm10708
258	24	0.9	186526	10	AL596083	Mouse DNA	c 331	23	0.9	591	3	AY241966	AY241966 Dermacent
259	24	0.9	188106	2	CR559942	Danio rer	c 332	23	0.9	592	11	BY065928	BY065928 S212P6917
260	24	0.9	188199	2	AC079242	Mouse muscu	c 333	23	0.9	593	11	BY047232	BY047232 S212P6917
261	24	0.9	192199	2	AC139312	Bos tauru	c 334	23	0.9	624	10	AY533231	AY533231 Chinchilli
262	24	0.9	194641	10	AC132462	Mouse muscu	c 335	23	0.9	632	3	AY193842	AY193842 Mastigamo
263	24	0.9	195905	2	CR392329	Danio rer	c 336	23	0.9	650	10	BC049564	BC049564 Mus muscu
264	24	0.9	198227	8	ATCRRIV43	Arabidops	c 337	23	0.9	726	6	BD021022	BD021022 Novel gen
265	24	0.9	199257	10	AL928545	Mouse DNA	c 338	23	0.9	726	6	BD100960	BD100960 Novel gen
266	24	0.9	199882	10	AC102145	Mouse muscu	c 339	23	0.9	746	9	BC064965	BC064965 Homo sapi
267	24	0.9	200538	10	AC134910	Mouse muscu	c 340	23	0.9	751	9	BC000606	BC000606 Homo sapi
268	24	0.9	200595	10	AC126670	Mouse muscu	c 341	23	0.9	765	8	AF220539	AF220539 Euphorbia
269	24	0.9	201630	2	BX629351	Danio rer	c 342	23	0.9	782	10	BC038931	BC038931 Mus muscu
270	24	0.9	201823	9	BX629351	Human DNA	c 343	23	0.9	799	6	BD021358	BD021358 Novel gen
271	24	0.9	204751	2	AC109220	Mouse muscu	c 344	23	0.9	799	6	BD101296	BD101296 Novel gen
272	24	0.9	205212	5	AL935173	Zebrafish	c 345	23	0.9	829	8	AY422285	AY422285 Lycopersi
273	24	0.9	205642	9	AC016586	Homo sapi	c 346	23	0.9	829	8	BD018606	BD018606 Novel gen
274	24	0.9	206486	10	AC136518	Mouse muscu	c 347	23	0.9	842	6	BD098544	BD098544 Novel gen
275	24	0.9	207318	10	AC125484	Mouse muscu	c 348	23	0.9	875	6	BOL508054	BOL508054 Braselica
276	24	0.9	209164	8	ATPCA5	Arabidops	c 349	23	0.9	876	5	AY539958	AY539958 Takifugu
277	24	0.9	209796	5	BX005407	Zebrafish	c 350	23	0.9	922	6	AX646757	AX646757 Sequence
278	24	0.9	210023	5	AC034116	Mouse muscu	c 351	23	0.9	972	9	BC002614	BC002614 Homo sapi
279	24	0.9	212604	10	AC115948	Mouse muscu	c 352	23	0.9	997	9	AX026296	AX026296 Homo sapi
280	24	0.9	212728	2	AC118257	Mouse muscu	c 353	23	0.9	1009	4	AY429909	AY429909 Oryctolag
281	24	0.9	213033	2	AC131482	Rattus no	c 354	23	0.9	1019	5	BC027418	BC027418 Mus muscu
282	24	0.9	216389	2	BX640518	Danio rer	c 355	23	0.9	1046	5	BC054269	BC054269 Xenopus l
283	24	0.9	216392	2	CR556722	Danio rer	c 356	23	0.9	1057	3	AR466595	AR466595 Aedes aeg
284	24	0.9	216677	10	AC125643	Rattus no	c 357	23	0.9	1066	3	AR544834	AR544834 Calliosobr
285	24	0.9	217375	2	AC126943	Mouse muscu	c 358	23	0.9	1078	9	BC015117	BC015117 Homo sapi
286	24	0.9	218754	2	AC094350	Rattus no	c 359	23	0.9	1146	5	AB093538	AB093538 Rana rugo
287	24	0.9	220970	2	AC137865	Mouse muscu	c 360	23	0.9	1147	5	AB093539	AB093539 Rana rugo
288	24	0.9	221928	5	BX120008	Zebrafish	c 361	23	0.9	1151	9	BT003978	BT003978 Arabidops
289	24	0.9	222072	2	AC141069	Homo sapi	c 362	23	0.9	1153	9	BC050998	BC050998 Homo sapi
290	24	0.9	222105	2	AC142172	Rattus no	c 363	23	0.9	1154	5	AB093537	AB093537 Rana rugo
291	24	0.9	224721	2	AC139632	Bos tauru	c 364	23	0.9	1162	6	FR118579	FR118579 Flavertia bl
292	24	0.9	224753	5	AL954663	Zebrafish	c 365	23	0.9	1185	6	AR452895	AR452895 Sequence
293	24	0.9	225612	10	AC137558	Mouse muscu	c 366	23	0.9	1194	8	AB107690	AB107690 Nicotiana
294	24	0.9	227298	10	AC099466	Rattus no	c 367	23	0.9	1240	8	AF054512	AF054512 Aspergill
295	24	0.9	228035	2	BX957332	Danio rer	c 368	23	0.9	1259	3	U35314	U35314 Penaeus van
296	24	0.9	228448	10	AL596456	Mouse DNA	c 369	23	0.9	1270	9	BC015088	BC015088 Homo sapi
297	24	0.9	228842	2	AE003720	Drosophila	c 370	23	0.9	1271	9	BC012543	BC012543 Homo sapi
298	24	0.9	233731	2	AC098558	Rattus no	c 371	23	0.9	1278	6	AX443077	AX443077 Sequence
299	24	0.9	236595	2	AC096444	Rattus no	c 372	23	0.9	1278	6	AX443079	AX443079 Sequence
300	24	0.9	238055	2	AC131139	Rattus no	c 373	23	0.9	1297	8	AB122122	AB122122 Porphyra
301	24	0.9	238075	2	AC127907	Rattus no	c 374	23	0.9	1300	9	BC013925	BC013925 Homo sapi
302	24	0.9	238965	2	AC126647	Rattus no	c 375	23	0.9	1310	5	AB008457	AB008457 Rana rugo
303	24	0.9	240519	2	AC105698	Rattus no	c 376	23	0.9	1314	5	AB008456	AB008456 Daucus ca
304	24	0.9	241726	2	AC097121	Rattus no	c 377	23	0.9	1316	5	AB027706	AB027706 Daucus ca
305	24	0.9	242397	2	AC130971	Rattus no	c 378	23	0.9	1318	5	AB008461	AB008461 Rana rugo
306	24	0.9	245317	2	AC116210	Rattus no	c 379	23	0.9	1320	5	AY273813	AY273813 Coffea ca
307	24	0.9	250517	2	AC094734	Rattus no	c 380	23	0.9	1332	5	BC056581	BC056581 Danio rer
308	24	0.9	253924	3	AE014822	Plasmodu	c 381	23	0.9	1336	8	CSCHRCPT	CSCHRCPT C. ecalvus m
309	24	0.9	254314	2	AC122953	Rattus no	c 382	23	0.9	1340	8	AY102637	AY102637 Arabidops
310	24	0.9	256759	2	AC126002	Rattus no	c 383	23	0.9	1344	8	MAN251920	MAN251920 Metarhizi
311	24	0.9	257167	2	AC134608	Mouse muscu	c 384	23	0.9	1346	9	HSU79286	HSU79286 Human argin

385	23	0.9	1347	6	AR150709	Sequence	458	23	0.9	3642	3	AY060636	AY060636 Drosophila
386	23	0.9	1369	8	AY443006	Amma maju	459	23	0.9	4500	9	HSMB07272	BS647128 Homo sapi
387	23	0.9	1380	8	AY325817	AY325817 Lycopersi	460	23	0.9	4500	9	HSMB07272	BS647128 Homo sapi
388	23	0.9	1388	5	BC075769	Danio rer	461	23	0.9	5243	6	AX528002	AX528002 Sequence
389	23	0.9	1395	9	BC008043	Homo sapi	462	23	0.9	5243	6	AX528002	AX528002 Sequence
390	23	0.9	1434	3	AF166600	AF166600 Dicycoste	463	23	0.9	5397	9	AF1538182	AF1538182 Homo sapi
391	23	0.9	1480	9	BC009827	BC009827 Homo sapi	464	23	0.9	5398	3	PI1439609	PI1439609 Paracentr
392	23	0.9	1481	9	AF055467	AF055467 Homo sapi	465	23	0.9	8731	4	AT293846	AT293846 Canis fam
393	23	0.9	1502	8	BT014141	BT014141 Lycopersi	466	23	0.9	8800	6	BD142516	BD142516 Med.ificat
394	23	0.9	1537	6	CQ493203	CQ493203 Sequence	467	23	0.9	8800	6	BD143100	BD143100 Method of
395	23	0.9	1543	3	HAU41810	HAU41810 Homarus ame	468	23	0.9	20809	9	AL359255	AL359255 Human DNA
396	23	0.9	1563	8	BT014372	BT014372 Lycopersi	469	23	0.9	27243	10	BX000354	BX000354 Mouse DNA
397	23	0.9	1595	10	BC010827	BC010827 Mus muscu	470	23	0.9	27856	8	AB090951	AB090951 Arabidops
398	23	0.9	1665	8	AY162466	AY162466 Spiraeta n	471	23	0.9	28486	6	AX695350	AX695350 Sequence
399	23	0.9	1700	9	BC050426	BC050426 Homo sapi	472	23	0.9	28790	10	BX517181	BX517181 Mouse DNA
400	23	0.9	1737	10	BC004037	BC004037 Mus muscu	473	23	0.9	35229	8	AC015986	AC015986 Arabidops
401	23	0.9	1835	3	AY061554	AY061554 Drosophila	474	23	0.9	38483	8	ATY12776	ATY12776 Arabidops
402	23	0.9	1880	5	BC073563	BC073563 Xenopus l	475	23	0.9	38592	8	EX119958	EX119958 Human DNA
403	23	0.9	1891	9	AF131839	AF131839 Homo sapi	476	23	0.9	39688	2	CR478038	CR478038 Mus muscu
404	23	0.9	1895	4	AF076180	AF076180 Trichostur	477	23	0.9	42801	8	AC079677	AC079677 Arabidops
405	23	0.9	1952	9	BC007508	BC007508 Homo sapi	478	23	0.9	44560	9	AC004536	AC004536 Homo sapi
406	23	0.9	1952	9	BC053666	BC053666 Homo sapi	479	23	0.9	45813	9	AC007165	AC007165 Homo sapi
407	23	0.9	1959	9	AK025734	AK025734 Homo sapi	480	23	0.9	48128	6	AX059512	AX059512 Sequence
408	23	0.9	1975	5	BC054966	BC054966 Xenopus l	481	23	0.9	48891	8	AC082643	AC082643 Arabidops
409	23	0.9	1976	5	AY333289	AY333289 Petromyco	482	23	0.9	52094	2	TI61118	TI61118 Arabidops
410	23	0.9	1982	9	BC007439	BC007439 Homo sapi	483	23	0.9	52717	8	AB015227	AB015227 Arabidops
411	23	0.9	2000	6	AX461222	AX461222 Sequence	484	23	0.9	53095	2	AC150705	AC150705 Medicago
412	23	0.9	2000	6	AX508028	AX508028 Sequence	485	23	0.9	53637	2	BS000217	BS000217 Pan trogl
413	23	0.9	2000	6	AX508396	AX508396 Sequence	486	23	0.9	54528	8	AB020754	AB020754 Arabidops
414	23	0.9	2000	6	AX510422	AX510422 Sequence	487	23	0.9	54987	8	AB026632	AB026632 Arabidops
415	23	0.9	2033	8	SPU63635	SPU63635 Schistosom	488	23	0.9	56942	2	AC100571	AC100571 Mus muscu
416	23	0.9	2036	6	BD186962	BD186962 Nucleic a	489	23	0.9	57985	2	AC100213	AC100213 Mus muscu
417	23	0.9	2062	3	AY499540	AY499540 Aedes aeg	490	23	0.9	58780	2	AC005085	AC005085 Homo sapi
418	23	0.9	2078	9	HS4275986	AJ2715986 Homo sapi	491	23	0.9	61510	8	AB028606	AB028606 Arabidops
419	23	0.9	2081	5	BC071523	BC071523 Danio rer	492	23	0.9	65073	2	AC119182	AC119182 Mus muscu
420	23	0.9	2142	3	BC077749	BC077749 Xenopus l	493	23	0.9	66237	8	FC214	FC214 Arabidops
421	23	0.9	2186	3	AY058696	AY058696 Drosophila	494	23	0.9	66237	8	AC007266	AC007266 Arabidops
422	23	0.9	2202	9	AB050508	AB050508 Macaca fa	495	23	0.9	66796	2	AC017435	AC017435 Drosophila
423	23	0.9	2263	9	AK026764	AK026764 Homo sapi	496	23	0.9	67322	2	AC102007	AC102007 Mus muscu
424	23	0.9	2288	3	CIN251957	AJ261957 Clona int	497	23	0.9	69602	2	AC101382	AC101382 Mus muscu
425	23	0.9	2300	9	HSMB04093	AL832782 Homo sapi	498	23	0.9	71606	2	AC149806	AC149806 Medicago
426	23	0.9	2338	6	AX857442	AX857442 Sequence	499	23	0.9	72718	2	AC016268	AC016268 Homo sapi
427	23	0.9	2338	6	BD155735	BD155735 Primer fo	500	23	0.9	73635	8	AC011915	AC011915 Arabidops
428	23	0.9	2338	9	AK000952	AK000952 Homo sapi	501	23	0.9	73920	10	AC091397	AC091397 Rattus no
429	23	0.9	2345	5	BC078346	BC078346 Danio rer	502	23	0.9	74026	2	AC101196	AC101196 Mus muscu
430	23	0.9	2349	10	BC004739	BC004739 Mus muscu	503	23	0.9	74494	8	AC007171	AC007171 Arabidops
431	23	0.9	2405	6	AR050135	AR050135 Sequence	504	23	0.9	74999	9	AC008960	AC008960 Homo sapi
432	23	0.9	2405	6	AR050142	AR050142 Sequence	505	23	0.9	75519	2	AC124999	AC124999 Mus muscu
433	23	0.9	2500	9	BC017040	BC017040 Homo sapi	506	23	0.9	77290	8	AC074284	AC074284 Arabidops
434	23	0.9	2540	10	BC074005	BC074005 Rattus no	507	23	0.9	77636	2	AB077822	AB077822 Arabidops
435	23	0.9	2545	9	HSMB02590	AL162059 Homo sapi	508	23	0.9	78514	2	AL356496	AL356496 Homo sapi
436	23	0.9	2552	10	AY102701	AY102701 Mus muscu	509	23	0.9	80309	8	AC022288	AC022288 Arabidops
437	23	0.9	2594	10	BC056435	BC056435 Mus muscu	510	23	0.9	80367	8	AC012329	AC012329 Arabidops
438	23	0.9	2629	9	HSMB06952	BX649183 Homo sapi	511	23	0.9	80374	8	TXK14	TXK14 Arabidops
439	23	0.9	2682	6	BD237779	BD237779 Apomixis	512	23	0.9	80413	8	AC021665	AC021665 Arabidops
440	23	0.9	2682	6	AX024562	AX024562 Sequence	513	23	0.9	80497	8	HS27K14	HS27K14 Human DNA
441	23	0.9	2715	10	MMU238213	AJ238213 Mus muscu	514	23	0.9	82356	8	AC006135	AC006135 Arabidops
442	23	0.9	2741	5	AY079197	AY079197 Raja egia	515	23	0.9	83689	8	AB017067	AB017067 Arabidops
443	23	0.9	2839	9	AC013018	AC013018 Drosophila	516	23	0.9	83796	8	AC004165	AC004165 Arabidops
444	23	0.9	2839	9	AC019455	AC019455 Drosophila	517	23	0.9	84203	8	AC005106	AC005106 Genomic s
445	23	0.9	2847	10	BC006671	BC006671 Mus muscu	518	23	0.9	84687	5	BX510368	BX510368 Zebrafish
446	23	0.9	2888	9	BC022050	BC022050 Homo sapi	519	23	0.9	84785	8	AB002064	AB002064 Arabidops
447	23	0.9	2910	6	AK431097	AK431097 Sequence	520	23	0.9	85347	2	AC017560	AC017560 Drosophila
448	23	0.9	2998	3	AY051780	AY051780 Drosophila	521	23	0.9	85567	2	AB016878	AB016878 Arabidops
449	23	0.9	3004	8	BT014535	BT014535 Lycopersi	522	23	0.9	85718	10	AL806530	AL806530 Mouse DNA
450	23	0.9	3030	5	BC063345	BC063345 Xenopus t	523	23	0.9	86001	8	AB016872	AB016872 Arabidops
451	23	0.9	3112	10	AY245772	AY245772 Monosiga	524	23	0.9	86801	8	AC007190	AC007190 Genomic s
452	23	0.9	3112	10	BC025944	BC025944 Mus muscu	525	23	0.9	86826	3	PFMA13P5	PFMA13P5 Arabidops
453	23	0.9	3189	3	AK115035	AK115035 Clona int	526	23	0.9	87630	8	AB016889	AB016889 Arabidops
454	23	0.9	3292	9	HSMB07557	BX647412 Homo sapi	527	23	0.9	87951	10	AL954855	AL954855 Mouse DNA
455	23	0.9	3310	5	BC076848	BC076848 Xenopus l	528	23	0.9	90551	8	ATT5K18	ATT5K18 Arabidops
456	23	0.9	3558	9	HSMB03467	AL832160 Homo sapi	529	23	0.9	90823	8	ATT209	ATT209 Arabidops
457	23	0.9	3609	8	AB102956	AB102956 Zea mays	530	23	0.9	91106	8	AC006438	AC006438 Arabidops



C 531	23	0.9	92569	8	AC007166	AC007166 Arabidops	604	23	0.9	117854	9	HSJ85A10	AL117363 Human DNA
C 532	23	0.9	92898	2	ATP18009	AT138009 Dario rer	C 605	23	0.9	117930	8	ATP7J8	AC073271 Homo sapi
C 533	23	0.9	93038	2	ATP8P16	AL021633 Arabidops	C 606	23	0.9	118507	8	ATP7J8	AL137168 Arabidops
C 534	23	0.9	94016	2	AC148818	AC148818 Arabidops	C 607	23	0.9	118718	8	ATP18A5	AL035528 Arabidops
C 535	23	0.9	94635	1	AC139931	AC139931 Homo sapi	C 608	23	0.9	118767	10	BX000344	BX000344 Mouse DNA
C 536	23	0.9	95111	8	ATF27K19	AL163832 Arabidops	C 609	23	0.9	118829	8	CRS58722	CRS58722 M. truncat
C 537	23	0.9	95559	8	AC123570	AC123570 Medicago	C 610	23	0.9	119748	10	AL691430	AL691430 Mouse DNA
C 538	23	0.9	95896	9	AL356384	AL356384 Human DNA	C 611	23	0.9	120721	2	AC149495	AC149495 Medicago
C 539	23	0.9	95959	8	AC006217	AC006217 Arabidops	C 612	23	0.9	121701	10	AL644484	AL644484 Mouse DNA
C 540	23	0.9	96808	8	AC011914	AC011914 Arabidops	C 613	23	0.9	121884	8	AC123975	AC123975 Medicago
C 541	23	0.9	97451	8	ATF21E1	AL391716 Arabidops	C 614	23	0.9	121909	8	HSJ492P14	AL121977 Human DNA
C 542	23	0.9	97792	2	AC148763	AC148763 Medicago	C 615	23	0.9	122013	3	AC092399	AC092399 Drosophila
C 543	23	0.9	98552	2	AC150778	AC150778 Medicago	C 616	23	0.9	122517	10	AL645546	AL645546 Mouse DNA
C 544	23	0.9	98852	10	AL935127	AL935127 Mouse DNA	C 617	23	0.9	122538	10	AL844557	AL844557 Mouse DNA
C 545	23	0.9	98852	5	BX004777	BX004777 Zebrafish	C 618	23	0.9	123521	9	HSJ75G13	AL006998 Homo sapi
C 546	23	0.9	99804	8	AC006951	AC006951 Arabidops	C 619	23	0.9	123800	8	AC144541	AC144541 Medicago
C 547	23	0.9	100057	10	AL929178	AL929178 Mouse DNA	C 620	23	0.9	124254	10	AP450245	AP450245 Mus muscu
C 548	23	0.9	100228	9	AL451130	AL451130 Human DNA	C 621	23	0.9	124828	8	AC140913	AC140913 Medicago
C 549	23	0.9	100269	8	ATP18O22	AL163817 Arabidops	C 622	23	0.9	125204	2	AC115066	AC115066 Mus muscu
C 550	23	0.9	100285	10	AL929036	AL929036 Mouse DNA	C 623	23	0.9	125478	5	BX548016	BX548016 Zebrafish
C 551	23	0.9	100685	8	AC023754	AC023754 Arabidops	C 624	23	0.9	126915	2	AC127674	AC127674 Medicago
C 552	23	0.9	100906	8	ATP24G16	AL136647 Arabidops	C 625	23	0.9	126954	9	AL365440	AL365440 Human DNA
C 553	23	0.9	102004	8	AC005395	AC005395 Arabidops	C 626	23	0.9	127163	2	AC122730	AC122730 Medicago
C 554	23	0.9	102331	8	AC007120	AC007120 Arabidops	C 627	23	0.9	127277	9	AC010175	AC010175 Homo sapi
C 555	23	0.9	103517	8	AC073395	AC073395 Arabidops	C 628	23	0.9	127387	2	BX324148	BX324148 Dario rer
C 556	23	0.9	104204	8	ATP9C5	AL133964 Arabidops	C 629	23	0.9	127643	5	BX321877	BX321877 Zebrafish
C 557	23	0.9	105273	9	AL671884	AL671884 Human DNA	C 630	23	0.9	128501	9	AC024568	AC024568 Homo sapi
C 558	23	0.9	105294	8	AC126792	AC126792 Medicago	C 631	23	0.9	128533	2	AC146865	AC146865 Medicago
C 559	23	0.9	105335	2	AC131550	AC131550 Rattus no	C 632	23	0.9	128836	2	AC015893	AC015893 Mus muscu
C 560	23	0.9	105607	2	AC149582	AC149582 Medicago	C 633	23	0.9	129273	2	AC142018	AC142018 Rattus no
C 561	23	0.9	106454	8	AC141864	AC141864 Medicago	C 634	23	0.9	129328	10	AL935275	AL935275 Mouse DNA
C 562	23	0.9	106643	9	AC114783	AC114783 Homo sapi	C 635	23	0.9	130054	8	AC133862	AC133862 Medicago
C 563	23	0.9	106702	8	ATP7K2	AL033545 Arabidops	C 636	23	0.9	130223	2	AC107684	AC107684 Mus muscu
C 564	23	0.9	106753	8	AC126007	AC126007 Medicago	C 637	23	0.9	131691	9	AC096765	AC096765 Homo sapi
C 565	23	0.9	106763	9	AP002091	AP002091 Homo sapi	C 638	23	0.9	131732	10	AP259071	AP259071 Mus muscu
C 566	23	0.9	106973	9	U89959	U89959 Arabidops	C 639	23	0.9	132293	2	AC108187	AC108187 Felis cat
C 567	23	0.9	107137	9	AC008493	AC008493 Homo sapi	C 640	23	0.9	133159	10	AL928771	AL928771 Mouse DNA
C 568	23	0.9	107234	8	F24B9	AC007583 Arabidops	C 641	23	0.9	133379	8	AC083835	AC083835 Arabidops
C 569	23	0.9	108459	9	ATP72A6	AL078637 Arabidops	C 642	23	0.9	134043	2	AC146717	AC146717 Rattus no
C 570	23	0.9	109067	9	AC010460	AC010460 Homo sapi	C 643	23	0.9	135837	2	AC004837	AC004837 Homo sapi
C 571	23	0.9	109092	10	AL928803	AL928803 Mouse DNA	C 644	23	0.9	136786	2	AC124778	AC124778 Mus muscu
C 572	23	0.9	109180	8	AC064879	AC064879 Arabidops	C 645	23	0.9	137519	8	ATP1P17	AL049730 Arabidops
C 573	23	0.9	109891	9	HS112K5	Z85987 Human DNA	C 646	23	0.9	138144	2	AL928864	AL928864 Homo sapi
C 574	23	0.9	110000	2	AC106698	Continuaction (2 of	C 647	23	0.9	138271	9	AC006360	AC006360 Homo sapi
C 575	23	0.9	110000	2	AC110950	Continuaction (4 of	C 648	23	0.9	139193	2	BX004801	BX004801 Dario rer
C 576	23	0.9	110000	2	AC114902	Continuaction (2 of	C 649	23	0.9	139377	9	AC026692	AC026692 Homo sapi
C 577	23	0.9	110000	2	BX569799	Continuaction (4 of	C 650	23	0.9	139799	2	CR450693	CR450693 Dario rer
C 578	23	0.9	110000	2	PFMAL13_07	Continuaction (8 of	C 651	23	0.9	140624	8	AC124957	AC124957 Medicago
C 579	23	0.9	110000	2	PFMAL13_15	Continuaction (16 of	C 652	23	0.9	141228	2	BX511186	BX511186 Dario rer
C 580	23	0.9	110000	2	PFMAL7P1_05	Continuaction (7 of	C 653	23	0.9	141373	2	AC148528	AC148528 Medicago
C 581	23	0.9	110000	2	PFMAL7P1_09	Continuaction (10 of	C 654	23	0.9	141854	2	AC021555	AC021555 Homo sapi
C 582	23	0.9	110000	3	AC116305	AC116305 Dictyoste	C 655	23	0.9	142333	10	AC131798	AC131798 Homo sapi
C 583	23	0.9	110572	8	AE008683_2	Continuaction (3 of	C 656	23	0.9	142978	9	AC097717	AC097717 Homo sapi
C 584	23	0.9	111037	2	AC090067	AC147741 Medicago	C 657	23	0.9	143722	10	AL731792	AL731792 Mouse DNA
C 585	23	0.9	111037	2	AC090067	AL954694 Zebrafish	C 658	23	0.9	143728	10	AL645476	AL645476 Mouse DNA
C 586	23	0.9	111945	5	AL954694	AL954694 Zebrafish	C 659	23	0.9	143864	2	AC048359	AC048359 Homo sapi
C 587	23	0.9	111820	2	CR387919	CR387919 Dario rer	C 660	23	0.9	144217	8	AC139748	AC139748 Medicago
C 588	23	0.9	111945	2	ATP1C12	AL022224 Arabidops	C 661	23	0.9	144319	9	AC112218	AC112218 Homo sapi
C 589	23	0.9	112032	2	AC145221	AL049656 Arabidops	C 662	23	0.9	144533	10	AC121938	AC121938 Mus muscu
C 590	23	0.9	112098	2	AC144513	AC144513 Medicago	C 663	23	0.9	144659	9	AC006145	AC006145 Homo sapi
C 591	23	0.9	112154	8	AC146342	AC146342 Medicago	C 664	23	0.9	144734	2	AC148922	AC148922 Dasyu n
C 592	23	0.9	112369	2	AC005861	AC005861 Arabidops	C 665	23	0.9	144803	2	AC079555	AC079555 Mus muscu
C 593	23	0.9	112846	8	AP006113	AP006113 Lotus cor	C 666	23	0.9	144803	9	AC108101	AC108101 Homo sapi
C 594	23	0.9	112917	8	AC123976	AC123976 Medicago	C 667	23	0.9	145019	2	AC053488	AC053488 Homo sapi
C 595	23	0.9	113345	8	ATF6G15	AL049656 Arabidops	C 668	23	0.9	145443	10	AL645824	AL645824 Mouse DNA
C 596	23	0.9	113970	8	ATF6G15	AC096742 Homo sapi	C 669	23	0.9	145786	2	AC132929	AC132929 Mus muscu
C 597	23	0.9	114418	8	AC0067047	AC0067047 Arabidops	C 670	23	0.9	146061	10	AL627348	AL627348 Sequence
C 598	23	0.9	114449	8	AC146790	AC146790 Medicago	C 671	23	0.9	146505	9	F21M12	AC016194 Homo sapi
C 599	23	0.9	115365	2	AC149197	AC149197 Medicago	C 672	23	0.9	146768	9	AC016194	AC016194 Homo sapi
C 600	23	0.9	115622	2	AC148349	AC148349 Medicago	C 673	23	0.9	147103	2	CRS47131	CRS47131 Dario rer
C 601	23	0.9	116202	2	AC148349	AC148349 Medicago	C 674	23	0.9	147103	5	AL808141	AL808141 Zebrafish
C 602	23	0.9	116265	5	AC140941	AC140941 Gallus ga	C 675	23	0.9	147290	10	AC099608	AC099608 Mus muscu
C 603	23	0.9	116425	2	AC148395	AC148395 Medicago	C 676	23	0.9	149288	2	CR376830	CR376830 Dario rer

c 677	23	0.9 149391	9	AC093825	AC093825 Homo sapi	750	23	0.9 166004	9	AL935036	AL935036 Human DNA
c 678	23	0.9 149442	5	BX255899	BX255899 Zebrafish	c 751	23	0.9 166110	5	AL929305	AL929305 Zebrafish
c 679	23	0.9 149603	5	BX248414	BX248414 Zebrafish	752	23	0.9 166114	5	AL807829	AL807829 Zebrafish
c 680	23	0.9 150392	2	AC004921	AC004921 Homo sapi	753	23	0.9 166518	9	HS120022	AL031847 Human DNA
c 681	23	0.9 150604	2	AC142030	AC142030 Rattus no	c 754	23	0.9 166555	10	AC124366	AC124366 Mus muscu
c 682	23	0.9 150756	2	AC130793	AC130793 Canis fam	c 755	23	0.9 166928	2	AC110905	AC110905 Mus muscu
c 683	23	0.9 150789	9	HS21F7	AL033375 Human DNA	c 756	23	0.9 166978	2	AC004098	AC004098 Homo sapi
c 684	23	0.9 151432	2	AC120398	AC120398 Mus muscu	c 757	23	0.9 167130	2	AC018390	AC018390 Homo sapi
c 685	23	0.9 151544	2	AC013509	AC013509 Homo sapi	c 758	23	0.9 167166	2	AC024654	AC024654 Homo sapi
c 686	23	0.9 152416	9	AC096663	AC096663 Homo sapi	759	23	0.9 167195	2	AC007808	AC007808 Drosophill
c 687	23	0.9 153736	9	AC129098	AC129098 Pan trogl	760	23	0.9 167310	10	AC105169	AC105169 Mus muscu
c 688	23	0.9 153746	9	AC108125	AC108125 Homo sapi	761	23	0.9 168067	2	BX914202	BX914202 Dantio rer
c 689	23	0.9 153751	3	AC116551	AC116551 Dictyoste	c 762	23	0.9 168077	2	AC114901	AC114901 Mus muscu
c 690	23	0.9 153959	2	AC144526	AC144526 Homo sapi	c 763	23	0.9 168353	2	AC146598	AC146598 Mus muscu
c 691	23	0.9 154108	5	BX005289	BX005289 Zebrafish	c 764	23	0.9 168425	2	BX908778	BX908778 Dantio rer
c 692	23	0.9 154480	10	AF834675	AF834675 Mus muscu	765	23	0.9 168504	3	AC115484	AC115484 Drosophill
c 693	23	0.9 154541	2	AC120212	AC120212 Canis fam	c 766	23	0.9 168814	10	AL591582	AL591582 Mouse DNA
c 694	23	0.9 154590	2	AC105216	AC105216 Homo sapi	c 767	23	0.9 168835	9	AL954210	AL954210 Pan trogl
c 695	23	0.9 154604	5	AL954739	AL954739 Zebrafish	c 768	23	0.9 169849	9	AC078988	AC078988 Homo sapi
c 696	23	0.9 154673	2	BX927192	BX927192 Dantio rer	c 769	23	0.9 170084	2	BX321918	BX321918 Dantio rer
c 697	23	0.9 155332	9	AC109581	AC109581 Homo sapi	c 770	23	0.9 170178	5	AL954848	AL954848 Zebrafish
c 698	23	0.9 155372	2	AC108818	AC108818 Mus muscu	c 771	23	0.9 170393	2	AC024304	AC024304 Homo sapi
c 699	23	0.9 155587	10	AC131788	AC131788 Mus muscu	c 772	23	0.9 170429	10	AC142146	AC142146 Mus muscu
c 700	23	0.9 155771	10	AC121909	AC121909 Mus muscu	c 773	23	0.9 170472	2	BX957303	BX957303 Dantio rer
c 701	23	0.9 156475	10	AC147380	AC147380 Mus muscu	c 774	23	0.9 170654	10	AC116502	AC116502 Mus muscu
c 702	23	0.9 156698	10	AL837509	AL837509 Mouse DNA	c 775	23	0.9 170655	2	AC105986	AC105986 Mus muscu
c 703	23	0.9 156699	2	CR318633	CR318633 Dantio rer	c 776	23	0.9 170658	2	AC004072	AC004072 Human Chr
c 704	23	0.9 156807	2	CR394547	CR394547 Dantio rer	777	23	0.9 170745	2	AC120835	AC120835 Rattus no
c 705	23	0.9 157162	9	AL360224	AL360224 Human DNA	c 778	23	0.9 170754	2	AC012403	AC012403 Homo sapi
c 706	23	0.9 157875	2	AC131893	AC131893 Gallus ga	c 779	23	0.9 170928	2	AC034148	AC034148 Homo sapi
c 707	23	0.9 158060	2	AC148206	AC148206 Callicebu	c 780	23	0.9 171175	9	AC135507	AC135507 Homo sapi
c 708	23	0.9 158096	8	AC007887	AC007887 Genomic s	781	23	0.9 171297	2	AC137743	AC137743 Mus muscu
c 709	23	0.9 158147	10	AC127280	AC127280 Mus muscu	782	23	0.9 171556	5	AL935284	AL935284 Zebrafish
c 710	23	0.9 158548	3	PFMAL3P2	AL034558 Plasmodiu	783	23	0.9 172126	10	AC122461	AC122461 Mus muscu
c 711	23	0.9 158617	9	AC135972	AC135972 Homo sapi	c 784	23	0.9 172954	5	AL929152	AL929152 Zebrafish
c 712	23	0.9 158781	2	AC024978	AC024978 Homo sapi	785	23	0.9 173084	2	BX000460	BX000460 Zebrafish
c 713	23	0.9 158980	9	HS046P5	AL080315 Human DNA	c 786	23	0.9 173370	10	AC144629	AC144629 Mus muscu
c 714	23	0.9 159254	2	BX897733	BX897733 Dantio rer	c 787	23	0.9 173602	9	AC090226	AC090226 Homo sapi
c 715	23	0.9 159369	10	AL929499	AL929499 Mouse DNA	788	23	0.9 173617	5	BX323875	BX323875 Zebrafish
c 716	23	0.9 159377	2	AC123749	AC123749 Mus muscu	789	23	0.9 173624	9	AC150657	AC150657 Pan trogl
c 717	23	0.9 159734	10	AC122533	AC122533 Mus muscu	790	23	0.9 173801	2	AC129293	AC129293 Mus muscu
c 718	23	0.9 159989	2	AC138687	AC138687 Homo sapi	791	23	0.9 174358	2	CR536514	CR536514 Dantio rer
c 719	23	0.9 160004	9	AL139318	AL139318 Human DNA	792	23	0.9 174397	9	AC067773	AC067773 Homo sapi
c 720	23	0.9 160145	9	AL161612	AL161612 Human DNA	c 793	23	0.9 174411	10	AC116120	AC116120 Mus muscu
c 721	23	0.9 160246	2	AC0099221	AC0099221 Homo sapi	794	23	0.9 174471	10	AC122534	AC122534 Mus muscu
c 722	23	0.9 160262	9	AC005242	AC005242 Homo sapi	795	23	0.9 174869	2	AP001840	AP001840 Homo sapi
c 723	23	0.9 160390	2	AC068186	AC068186 Homo sapi	c 796	23	0.9 175187	5	BX511240	BX511240 Zebrafish
c 724	23	0.9 160571	10	AL672008	AL672008 Ootolemur	c 797	23	0.9 175346	2	AC115659	AC115659 Rattus no
c 725	23	0.9 160602	2	AC148622	AC148622 Homo sapi	c 798	23	0.9 176245	10	AC124580	AC124580 Mus muscu
c 726	23	0.9 160736	2	AC027477	AC027477 Homo sapi	c 799	23	0.9 176262	2	AC118273	AC118273 Mus muscu
c 727	23	0.9 161081	2	AC022997	AC022997 Homo sapi	c 800	23	0.9 176427	2	AC102743	AC102743 Mus muscu
c 728	23	0.9 161318	9	AB045359	AB045359 Homo sapi	c 801	23	0.9 176893	5	AL928685	AL928685 Zebrafish
c 729	23	0.9 161474	5	AC144356	AC144356 Dantio rer	c 802	23	0.9 176929	2	AC004887	AC004887 Homo sapi
c 730	23	0.9 161644	5	BX294104	BX294104 Zebrafish	c 803	23	0.9 177203	2	AC055784	AC055784 Homo sapi
c 731	23	0.9 161710	10	AC128736	AC128736 Mus muscu	c 804	23	0.9 177464	2	AC069221	AC069221 Homo sapi
c 732	23	0.9 162054	2	AC046201	AC046201 Homo sapi	805	23	0.9 177528	2	AC093444	AC093444 Pan trogl
c 733	23	0.9 162139	9	AC147651	AC147651 Homo sapi	c 806	23	0.9 177675	10	AC124534	AC124534 Mus muscu
c 734	23	0.9 162394	10	AL683891	AL683891 Mouse DNA	807	23	0.9 177757	2	CR388378	CR388378 Dantio rer
c 735	23	0.9 162409	2	AC022184	AC022184 Homo sapi	c 808	23	0.9 177797	2	AC016581	AC016581 Homo sapi
c 736	23	0.9 163247	2	BX571697	BX571697 Dantio rer	809	23	0.9 178093	2	AC137848	AC137848 Mus muscu
c 737	23	0.9 163373	2	AC101896	AC101896 Mus muscu	810	23	0.9 178568	9	HS1164C1	HS1164C1 Human DNA
c 738	23	0.9 163676	5	AL845550	AL845550 Zebrafish	c 811	23	0.9 178596	9	AC097015	AC097015 Homo sapi
c 739	23	0.9 163788	5	BX571822	BX571822 Zebrafish	c 812	23	0.9 178717	2	BX649315	BX649315 Dantio rer
c 740	23	0.9 163810	2	AC116802	AC116802 Mus muscu	c 813	23	0.9 178783	2	AC079486	AC079486 Mus muscu
c 741	23	0.9 163982	2	CR381576	CR381576 Dantio rer	c 814	23	0.9 179224	2	AC122515	AC122515 Mus muscu
c 742	23	0.9 163982	5	AL954708	AL954708 Zebrafish	c 815	23	0.9 179845	2	AY178787	AY178787 Canis fam
c 743	23	0.9 164026	9	AC058791	AC058791 Homo sapi	c 816	23	0.9 179956	2	AC098465	AC098465 Rattus no
c 744	23	0.9 164030	10	AC125073	AC125073 Mus muscu	c 817	23	0.9 180137	2	BX957257	BX957257 Dantio rer
c 745	23	0.9 164232	2	AC067792	AC067792 Homo sapi	c 818	23	0.9 180279	2	BX901957	BX901957 Homo sapi
c 746	23	0.9 164611	2	AC128858	AC128858 Rattus no	c 819	23	0.9 180690	2	AC122469	AC122469 Mus muscu
c 747	23	0.9 164670	2	AC102297	AC102297 Mus muscu	c 820	23	0.9 180878	2	AC118641	AC118641 Homo sapi
c 748	23	0.9 164684	9	AL590683	AL590683 Human DNA	c 821	23	0.9 180948	9	AC114399	AC114399 Homo sapi
c 749	23	0.9 165910	9	AL928668	AL928668 Human DNA	c 822	23	0.9 181061	9	AC115282	AC115282 Homo sapi

C 823	23	0.9	181408	2	AC101891	AC101891 Mus muscu	C 896	23	0.9	196107	8	ATCRRIV34	AL161534 Arabidops
C 824	23	0.9	181430	2	CR407555	CR407555 Danio rer	C 897	23	0.9	196424	2	AC118884	AC118884 Rattus no
C 825	23	0.9	181467	10	AL844586	AL844586 Mouse DNA	C 898	23	0.9	196534	2	AC125895	AC125895 Rattus no
C 826	23	0.9	181585	9	AL139350	AL139350 Human DNA	C 899	23	0.9	196756	2	CR556636	CR556636 Danto rer
C 827	23	0.9	181700	10	AC131231	AC131231 Mus muscu	C 900	23	0.9	196835	2	AC102099	AC102099 Mus muscu
C 828	23	0.9	182577	5	BX640466	BX640466 Zebrafish	C 901	23	0.9	196835	2	AC102099	AC102099 Mus muscu
C 829	23	0.9	182591	10	AC127272	AC127272 Mus muscu	C 902	23	0.9	196835	2	AC102099	AC102099 Mus muscu
C 830	23	0.9	182725	2	AC102786	AC102786 Mus muscu	C 903	23	0.9	196937	5	BX469913	BX469913 Zebrafish
C 831	23	0.9	183205	2	AC133734	AC133734 Rattus no	C 904	23	0.9	197273	10	AC024607	AC024607 Mus muscu
C 832	23	0.9	183222	10	AC123620	AC123620 Mus muscu	C 905	23	0.9	197852	10	AL606509	AL606509 Mouse DNA
C 833	23	0.9	183322	10	AC112990	AC112990 Mus muscu	C 906	23	0.9	198026	2	AC124730	AC124730 Mus muscu
C 834	23	0.9	183328	2	AC101773	AC101773 Mus muscu	C 907	23	0.9	198092	2	AL772345	AL772345 Mouse DNA
C 835	23	0.9	183496	2	CR391938	CR391938 Danto rer	C 908	23	0.9	198354	8	ATAP22	Z99708 Arabidopsis
C 836	23	0.9	183537	10	AC140196	AC140196 Mus muscu	C 909	23	0.9	198354	8	ATAP22	Z99708 Arabidopsis
C 837	23	0.9	183553	2	AC102335	AC102335 Homo sapi	C 910	23	0.9	198402	8	ATCRRIV61	AL161561 Arabidops
C 838	23	0.9	184011	2	CR388410	CR388410 Danto rer	C 911	23	0.9	198418	2	BX548034	BX548034 Danto rer
C 839	23	0.9	184289	9	AL359878	AL359878 Human DNA	C 912	23	0.9	198427	8	ATCRRIV52	AL161552 Arabidops
C 840	23	0.9	184729	9	AC097521	AC097521 Homo sapi	C 913	23	0.9	198470	9	AC046170	AC046170 Homo sapi
C 841	23	0.9	184737	2	AC124566	AC124566 Mus muscu	C 914	23	0.9	198669	8	ATCRRIV78	AL161582 Arabidops
C 842	23	0.9	184927	2	AC137236	AC137236 Rattus no	C 915	23	0.9	198750	8	ATCRRIV85	AL161569 Arabidops
C 843	23	0.9	185395	2	AC118699	AC118699 Mus muscu	C 916	23	0.9	198780	8	ATCRRIV86	AL161590 Arabidops
C 844	23	0.9	185597	2	AC126554	AC126554 Mus muscu	C 917	23	0.9	198780	8	ATCRRIV86	AL161590 Arabidops
C 845	23	0.9	185763	2	AC109240	AC109240 Mus muscu	C 918	23	0.9	198928	10	AC122212	AC122212 Mus muscu
C 846	23	0.9	186533	2	CR589881	CR589881 Danto rer	C 919	23	0.9	199382	8	ATCRRIV50	AL161550 Arabidops
C 847	23	0.9	186758	2	AC127243	AC127243 Mus muscu	C 920	23	0.9	199577	8	ATCRRIV57	AL161557 Arabidops
C 848	23	0.9	186812	9	AC110291	AC110291 Homo sapi	C 921	23	0.9	199634	8	ATCRRIV36	AL161537 Arabidops
C 849	23	0.9	187004	9	AC023154	AC023154 Homo sapi	C 922	23	0.9	199667	8	ATCRRIV37	AL161537 Arabidops
C 850	23	0.9	187226	10	AL528544	AL528544 Mouse DNA	C 923	23	0.9	199885	10	AC147376	AC147376 Mus muscu
C 851	23	0.9	187568	9	AP002967	AP002967 Homo sapi	C 924	23	0.9	200009	2	AC126208	AC126208 Rattus no
C 852	23	0.9	188046	2	BX649493	BX649493 Danto rer	C 925	23	0.9	200157	2	CR388376	CR388376 Danto rer
C 853	23	0.9	188246	9	AC013565	AC013565 Homo sapi	C 926	23	0.9	200420	9	AC138645	AC138645 Homo sapi
C 854	23	0.9	188564	2	BX548170	BX548170 Danto rer	C 927	23	0.9	200605	10	AC129314	AC129314 Mus muscu
C 855	23	0.9	189082	9	AL731537	AL731537 Human DNA	C 928	23	0.9	200665	10	AL683852	AL683852 Mouse DNA
C 856	23	0.9	189627	2	AC148926	AC148926 Rattus no	C 929	23	0.9	200814	5	BX571825	BX571825 Zebrafish
C 857	23	0.9	189781	2	AC134797	AC134797 Mus muscu	C 930	23	0.9	202086	10	AC080015	AC080015 Mus muscu
C 858	23	0.9	189924	2	AC119933	AC119933 Mus muscu	C 931	23	0.9	202156	10	AC102229	AC102229 Mus muscu
C 859	23	0.9	190131	2	AC116201	AC116201 Rattus no	C 932	23	0.9	202322	2	AC117316	AC117316 Mus muscu
C 860	23	0.9	190138	10	AC119432	AC119432 Mus muscu	C 933	23	0.9	202725	9	AC092128	AC092128 Homo sapi
C 861	23	0.9	190149	2	AC122097	AC122097 Rattus no	C 934	23	0.9	202855	10	AL589737	AL589737 Mouse DNA
C 862	23	0.9	190150	2	AC119157	AC119157 Mus muscu	C 935	23	0.9	202882	10	AC114006	AC114006 Mus muscu
C 863	23	0.9	190401	2	AC021974	AC021974 Homo sapi	C 936	23	0.9	203618	2	AC091282	AC091282 Mus muscu
C 864	23	0.9	190424	2	BX957227	BX957227 Danto rer	C 937	23	0.9	204147	2	AC145056	AC145056 Danto rer
C 865	23	0.9	190452	2	AC115853	AC115853 Mus muscu	C 938	23	0.9	204147	2	AC145056	AC145056 Danto rer
C 866	23	0.9	190544	2	AC131072	AC131072 Mus muscu	C 939	23	0.9	204287	10	AC126442	AC126442 Mus muscu
C 867	23	0.9	190547	2	AC102381	AC102381 Mus muscu	C 940	23	0.9	204325	2	BX284609	BX284609 Mus muscu
C 868	23	0.9	190559	2	AC117209	AC117209 Mus muscu	C 941	23	0.9	204701	2	BX572103	BX572103 Danto rer
C 869	23	0.9	190620	2	AC101717	AC101717 Mus muscu	C 942	23	0.9	204856	2	AC119275	AC119275 Mus muscu
C 870	23	0.9	190788	10	AC100730	AC100730 Mus muscu	C 943	23	0.9	205552	2	AC127949	AC127949 Mus muscu
C 871	23	0.9	190904	2	AC138081	AC138081 Mus muscu	C 944	23	0.9	206110	9	AL654203	AL654203 Pan trogl
C 872	23	0.9	190998	5	BX649394	BX649394 Zebrafish	C 945	23	0.9	206420	8	ATAP21	Z99707 Arabidopsis
C 873	23	0.9	191081	2	AC023143	AC023143 Homo sapi	C 946	23	0.9	206593	2	BX548015	BX548015 Danto rer
C 874	23	0.9	191317	2	AC150827	AC150827 Callithrix	C 947	23	0.9	206641	10	AC120394	AC120394 Mus muscu
C 875	23	0.9	191358	10	AC086608	AC086608 Mus muscu	C 948	23	0.9	207097	10	AL670708	AL670708 Mouse DNA
C 876	23	0.9	191410	10	AC083814	AC083814 Mus muscu	C 949	23	0.9	207159	10	AC117624	AC117624 Mus muscu
C 877	23	0.9	191473	2	AC127028	AC127028 Mus muscu	C 950	23	0.9	207432	10	AL844532	AL844532 Mouse DNA
C 878	23	0.9	191527	2	AC121306	AC121306 Mus muscu	C 951	23	0.9	207486	10	AL683876	AL683876 Mouse DNA
C 879	23	0.9	191638	10	AL663062	AL663062 Mouse DNA	C 952	23	0.9	207576	9	AC013722	AC013722 Homo sapi
C 880	23	0.9	191909	2	AC148337	AC148337 Homo sapi	C 953	23	0.9	207622	2	BX322531	BX322531 Danto rer
C 881	23	0.9	192187	3	AC117072	AC117072 Dictyoste	C 954	23	0.9	207945	10	AC123945	AC123945 Mus muscu
C 882	23	0.9	192543	2	AC128537	AC128537 Rattus no	C 955	23	0.9	208007	2	AC096847	AC096847 Papio anu
C 883	23	0.9	192937	9	AC087565	AC087565 Homo sapi	C 956	23	0.9	208280	2	AC102785	AC102785 Mus muscu
C 884	23	0.9	193593	10	AC111008	AC111008 Mus muscu	C 957	23	0.9	208341	2	AC147783	AC147783 Mus muscu
C 885	23	0.9	194175	2	AC109614	AC109614 Mus muscu	C 958	23	0.9	208498	2	BX679663	BX679663 Danto rer
C 886	23	0.9	194261	10	AC121264	AC121264 Mus muscu	C 959	23	0.9	208939	10	AC138342	AC138342 Mus muscu
C 887	23	0.9	194307	2	BX629332	BX629332 Danto rer	C 960	23	0.9	209091	2	AC102707	AC102707 Mus muscu
C 888	23	0.9	194721	5	BX470232	BX470232 Zebrafish	C 961	23	0.9	210249	2	AC115795	AC115795 Mus muscu
C 889	23	0.9	194917	2	CR456628	CR456628 Danto rer	C 962	23	0.9	210474	9	AC148447	AC148447 Pan trogl
C 890	23	0.9	194920	8	ATCRRIV74	ATCRRIV74 Rattus no	C 963	23	0.9	210636	9	AC006443	AC006443 Homo sapi
C 891	23	0.9	195313	2	AC133328	AC133328 Rattus no	C 964	23	0.9	210702	5	BX119982	BX119982 Zebrafish
C 892	23	0.9	195375	10	AC122747	AC122747 Rattus no	C 965	23	0.9	210822	2	AC116818	AC116818 Mus muscu
C 893	23	0.9	195771	10	AC108819	AC108819 Mus muscu	C 966	23	0.9	211063	2	AC069254	AC069254 Homo sapi
C 894	23	0.9	195771	10	AC108819	AC108819 Mus muscu	C 967	23	0.9	211333	10	AC138175	AC138175 Mus muscu
C 895	23	0.9	195817	2	AC119891	AC119891 Mus muscu	C 968	23	0.9	211814	2	BX470150	BX470150 Danto rer

c 969	23	0.9	212049	5	EX005484	EX005484 Zebrafish
970	23	0.9	213473	10	AC134869	AC134869 Mus muscu
971	23	0.9	213580	2	AC109224	AC109224 Mus muscu
c 972	23	0.9	213817	2	AC051615	AC051615 Mus muscu
c 973	23	0.9	213884	2	AC147475	AC147475 Mus muscu
c 974	23	0.9	213933	2	AC128930	AC128930 Rattus no
c 975	23	0.9	214438	10	AC124546	AC124546 Mus muscu
c 976	23	0.9	215019	10	AC107703	AC107703 Mus muscu
c 977	23	0.9	215571	10	AC116115	AC116115 Mus muscu
c 978	23	0.9	215742	9	AC148958	AC148958 Ocolemur
c 979	23	0.9	216172	9	AC098973	AC098973 Homo sapi
c 980	23	0.9	216444	2	AC116858	AC116858 Mus muscu
c 981	23	0.9	216501	10	AC107795	AC107795 Mus muscu
c 982	23	0.9	216502	2	AC109509	AC109509 Mus muscu
c 983	23	0.9	216567	2	CR558305	CR558305 Danio rer
c 984	23	0.9	216885	2	CR382328	CR382328 Mus muscu
c 985	23	0.9	218074	2	AC103155	AC103155 Rattus no
c 986	23	0.9	218283	2	AC132382	AC132382 Mus muscu
c 987	23	0.9	218688	10	AC132452	AC132452 Mus muscu
c 988	23	0.9	220193	10	AC131713	AC131713 Mus muscu
c 989	23	0.9	220454	2	CR382335	CR382335 Danio rer
c 990	23	0.9	220558	2	CR522881	CR522881 Danio rer
c 991	23	0.9	220874	2	CR376848	CR376848 Danio rer
c 992	23	0.9	221190	2	BX571704	BX571704 Danio rer
c 993	23	0.9	221405	2	AC114252	AC114252 Rattus no
c 994	23	0.9	221649	2	AC131083	AC131083 Mus muscu
c 995	23	0.9	222291	2	AC020860	AC020860 Mus muscu
c 996	23	0.9	222314	2	AC112362	AC112362 Rattus no
c 997	23	0.9	222358	2	AC099385	AC099385 Rattus no
c 998	23	0.9	222370	2	AC105895	AC105895 Rattus no
c 999	23	0.9	222805	2	AC133501	AC133501 Mus muscu
c1000	23	0.9	223443	2	AC112801	AC112801 Rattus no

## ALIGNMENTS

RESULT 1	AF170724	2679 bp	mRNA	linear	PII 03-AUG-2000
LOCUS	AF170724				
DEFINITION	Homo sapiens cell cycle checkpoint protein CHFR mRNA, complete cds.				
ACCESSION	AF170724				
VERSION	AF170724.1	GI:9651169			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	Scolnick,D.M. and Halazonetis,T.D.				
AUTHORS	Chfr defines a mitotic stress checkpoint that delays entry into metaphase				
TITLE	Nature 406 (6794), 430-435 (2000)				
JOURNAL	2 (bases 1 to 2679)				
MEDLINE	20388685				
PubMed	10935642				
REFERENCE	Halazonetis,T.D. and Scolnick,D.M.				
AUTHORS	Submitted (21-JUL-1999) Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA				
JOURNAL	Location/Qualifiers				
FEATURES	1..2679				
source	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	91..2085				
	/note="contains FHA domain and ring finger"				
	/codon_start=1				
	/product="cell cycle checkpoint protein CHFR"				
	/protein_id="AA91084.1"				
	/db_xref="GI:9651170"				
	/translation="MERPEEGKQSPPOPMGRLLRLGAESEPHVLLRRKREWTIGRR				
	GDLSPSNKLVSGDHCRIVVDEKSGQVLTLEDTSIGTIVNKLKVVKKQTCLQRTGVD				

## ORIGIN

Query Match 100.0%; Score 2679; DB 9; Length 2679;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AAGATTGGGACGACGAGCGCGCAATGCTCTTGACAGCGCGCGCGCGCGCGGTTCCGG	60
DB	1	AMGATTTGGACGACGAGCGCGCAATGCTCTTGACAGCGCGCGCGCGCGCGGTTCCGG	60
QY	61	GTTTCGGCGCGGCGCGGATGTGAATCCCGATGAGCGCGCGCGCGCGCGGATGTG	120
DB	61	GTTTCGGCGCGGCGCGGATGTGAATCCCGATGAGCGCGCGCGCGCGGATGTG	120
QY	121	CG	180
DB	121	CG	180
QY	181	GTCTCTCTGAGAGAGCGGAGTGCATCGCGCGCGAGCGAGGTTGGAAGCTTCTTC	240
DB	181	GTCTCTCTGAGAGAGCGGAGTGCATCGCGCGCGAGCGAGGTTGGAAGCTTCTTC	240
QY	241	CCGACCAATTAATCTGCTCTGAGATGATGTAATTTGATGATGATGATGATGATGAT	300
DB	241	CCGACCAATTAATCTGCTCTGAGATGATGTAATTTGATGATGATGATGATGATGAT	300
QY	301	CAGGTGACATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	360
DB	301	CAGGTGACATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	360
QY	361	AAGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	420
DB	361	AAGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	420
QY	421	GACCGGAAACAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT	480
DB	421	GACCGGAAACAAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT	480
QY	481	CAAGATCTTGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	540
DB	481	CAAGATCTTGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	540
QY	541	GAAGTGAAGGCGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT	600
DB	541	GAAGTGAAGGCGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT	600
QY	601	TGCTTTGAGAACCAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT	660
DB	601	TGCTTTGAGAACCAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT	660
QY	661	TCTTCCAGGAGCGCTTCTGCGAGGCGAGGCGGATGATGATGATGATGATGATGAT	720
DB	661	TCTTCCAGGAGCGCTTCTGCGAGGCGAGGCGGATGATGATGATGATGATGATGAT	720
QY	721	GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
DB	721	GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	780
QY	781	GCTTCAAGTCTTCCAGACAGAAAGATGATGATGATGATGATGATGATGATGAT	840
DB	781	GCTTCAAGTCTTCCAGACAGAAAGATGATGATGATGATGATGATGATGATGAT	840
QY	841	GAGATTTGAGCGCGTGAAGAAATGAGAGAGATGAGAGATGAGAGATGAGAGAT	900

841 GAGGATTTGAGGCGCGTGAAGAAATGAGAGGATGGGACCTTGACCTGAACGGG 900  
QY 901 CAGTTTGTGGTGGCACAACCGCTAGAAATGCCAAACCGTCCAGAGAGCTCAAGCA 960  
Db 901 CAGTTTGTGGTGGCACAACCGCTAGAAATGCCAAACCGTCCAGAGAGCTCAAGCA 960  
QY 961 GCGGCTGGGAAAGCAGACAGATGAGAGAGCGCTGACATGATCATCTGCCAGAGCTG 1020  
Db 961 GCGGCTGGGAAAGCAGACAGATGAGAGAGCGCTGACATGATCATCTGCCAGAGCTG 1020  
QY 1021 CTGCAAGACTGCGTGAATTGGCAGCCCTGATGACACGTTCTGCGCGCTTGCTA 1080  
Db 1021 CTGCAAGACTGCGTGAATTGGCAGCCCTGATGACACGTTCTGCGCGCTTGCTA 1080  
QY 1081 GCGTGGATGAGAGCGCTGCTCTGTGTCTTAAGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 1081 GCGTGGATGAGAGCGCTGCTCTGTGTCTTAAGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 AAAAAACACATCTCAACAACCTGATGAAGCATCTCAAGCATCAAGAGAT 1200  
Db 1141 AAAAAACACATCTCAACAACCTGATGAAGCATCTCAAGCATCAAGAGAT 1200  
QY 1201 CCGAGTGAAGAGATGTGCAAAATGATGATGCCAGATTAATCACTCAAGACATGCTG 1260  
Db 1201 CCGAGTGAAGAGATGTGCAAAATGATGATGCCAGATTAATCACTCAAGACATGCTG 1260  
QY 1261 CAGCCCAAAGTCAAGGCGGCTTTTCTGATGAAGAGAGAGGAGTCAAGAGACTGCTGAG 1320  
Db 1261 CAGCCCAAAGTCAAGGCGGCTTTTCTGATGAAGAGAGAGGAGTCAAGAGACTGCTGAG 1320  
QY 1321 CTGTGAGAGGTTGACAGTGAAGTCTGACATTAAGCAGCATAGCTGCTGCTGCTGAG 1380  
Db 1321 CTGTGAGAGGTTGACAGTGAAGTCTGACATTAAGCAGCATAGCTGCTGCTGCTGAG 1380  
QY 1381 TGTCTTGAATGACAGAGGAGCGGCGCAGCTTCCCACTGCCAGCAACCGAGGCGAG 1440  
Db 1381 TGTCTTGAATGACAGAGGAGCGGCGCAGCTTCCCACTGCCAGCAACCGAGGCGAG 1440  
QY 1441 CCGAGAGCCCGCAGAGGCGGCTGAGGATGACCTCTCAAGTCCGCTGACAGACAGCA 1500  
Db 1441 CCGAGAGCCCGCAGAGGCGGCTGAGGATGACCTCTCAAGTCCGCTGACAGACAGCA 1500  
QY 1501 GTTCCAGATTAAGTGAAGGCTGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
Db 1501 GTTCCAGATTAAGTGAAGGCTGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
QY 1561 CCCATGCCCGAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
Db 1561 CCCATGCCCGAGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
QY 1621 GCGGCTGCTGAGGCTTTTCTGCACTGTGATCTGAGGAGCTGACCCGAGCTGCTAC 1680  
Db 1621 GCGGCTGCTGAGGCTTTTCTGCACTGTGATCTGAGGAGCTGACCCGAGCTGCTAC 1680  
QY 1681 GCGTGGCTGAGGCGGCTTTTGTGATCTGATCTGAGGAGGAGGAGGAGGAGGAG 1740  
Db 1681 GCGTGGCTGAGGCGGCTTTTGTGATCTGATCTGAGGAGGAGGAGGAGGAGGAG 1740  
QY 1741 AACCAACAGAGCTACAGAGCTGAGACATCTGAAATTAATCTGAGCAACAGAGGTTGACA 1800  
Db 1741 AACCAACAGAGCTACAGAGCTGAGACATCTGAAATTAATCTGAGCAACAGAGGTTGACA 1800  
QY 1801 TGGAAAAACATGTTGACCGAGAGCTGCTGCTTCCAGCGGAGAGTGTCTGCTGCTCT 1860  
Db 1801 TGGAAAAACATGTTGACCGAGAGCTGCTGCTTCCAGCGGAGAGTGTCTGCTGCTCT 1860  
QY 1861 GATTACAGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920  
Db 1861 GATTACAGAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920  
QY 1921 GAGCTGACCTATCAAGTATGCGAGAAACATTCCTGCTTCCAGTGGCAGTGGCCCTAACA 1980

Db 1921 GAGCTGACCTATCAAGTATGCGAGAAACATTCCTGCTTCCAGTGGCAGTGGCCCTAACA 1980  
QY 1981 TCCCGTCCGATGCTATCTGAGGCGGCTGATCTGCCGACCTCAGGTGAAGCTCACACGCG 2040  
Db 1981 TCCCGTCCGATGCTATCTGAGGCGGCTGATCTGCCGACCTCAGGTGAAGCTCACACGCG 2040  
QY 2041 ATGAATTAATCAATCATCTGTGAACAGACAGAGGTTCAAAAATTAAGCATTCAGAGGCGCT 2100  
Db 2041 ATGAATTAATCAATCATCTGTGAACAGACAGAGGTTCAAAAATTAAGCATTCAGAGGCGCT 2100  
QY 2101 GAGCAGCTTTCAGCATGAGAGGTTGAAGAGAGGCTGTTTAAATACAGAGCAAGCAG 2160  
Db 2101 GAGCAGCTTTCAGCATGAGAGGTTGAAGAGAGGCTGTTTAAATACAGAGCAAGCAG 2160  
QY 2161 TGAAGGTGTTTTCAGAGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220  
Db 2161 TGAAGGTGTTTTCAGAGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220  
QY 2221 GGTGACTCTTCTGTGAGCTTTTACCTCTGATGAGACCTCCCAAGAGCCCGGCGG 2280  
Db 2221 GGTGACTCTTCTGTGAGCTTTTACCTCTGATGAGACCTCCCAAGAGCCCGGCGG 2280  
QY 2281 CCGCAGCCCGCTCTCTGTGAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340  
Db 2281 CCGCAGCCCGCTCTCTGTGAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340  
QY 2341 CGAAGCTTTCGTGAACAGGCGGCTCCCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400  
Db 2341 CGAAGCTTTCGTGAACAGGCGGCTCCCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400  
QY 2401 ATTTTCCGAACTACAGTTAAAGCAGAACTGTGTTTCAAGAAAGTTCAAGGAGAG 2460  
Db 2401 ATTTTCCGAACTACAGTTAAAGCAGAACTGTGTTTCAAGAAAGTTCAAGGAGAG 2460  
QY 2461 GGCAGATTATCAAAAAACATGTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520  
Db 2461 GGCAGATTATCAAAAAACATGTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520  
QY 2521 CGTACCAATATCCCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580  
Db 2521 CGTACCAATATCCCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580  
QY 2581 GGTGGTGTCTATCTTCAATTAAGTTAAATGTCACAACTGTGACCAATAATATA 2640  
Db 2581 GGTGGTGTCTATCTTCAATTAAGTTAAATGTCACAACTGTGACCAATAATATA 2640  
QY 2641 ATTATTAATTAACAATTGACAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 2679  
Db 2641 ATTATTAATTAACAATTGACAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 2679

RESULT 2  
AX877194 2639 bp DNA linear PART 17-DEC-2003  
LOCUS AX877194  
DEFINITION Sequence 12099 from Patent EP1074617.  
ACCESSION AX877194  
VERSION AX877194.1 GI:40031930  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 Ota, T., Iwagaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primers for synthesizing full-length cDNA and their use  
Patent: BP 1074617-A 12099 07-PB-2001;  
JOURNAL Research Association for Biotechnology (JBP)  
FEATURES  
source location/Qualifiers  
1..2639  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"





Qy	2389	GCTGTTTGTGCATTTTCCGAACACACAGTAACCAAGTCGTTTCAGGAAAAGTT	2448
Db	2341	GCTCTTTGTACATTTTCCGAACACTACAGTTAAACCAAGTCGTTTCAGGAAAAGTT	2400
Qy	2449	TCAAGGAGAAAGGCGCAAGTTATCAAAAAACATGTTTCAGGAGAAAGGAGCATTAAGTTA	2508
Db	2401	TCAAGGAGAGAAAGGCGCAAGTTATCAAAAAACATGTTTCAGGAGAAAGGAGCATTAAGTTA	2460
Qy	2509	CAGCCTAAGAACCGTACCAATATTCCTGCTGTCGGGAAAACACAGCATTTTAATCAATTT	2568
Db	2461	CAGCCTAAGAACCGTACCAATATTCCTGCTGTCGGGAAAACACAGCATTTTAATCAATTT	2520
Qy	2569	TTTATTTTAAATAGGTTTGTCGCTTAATCTTCAATAGAATTTAAATATGTCACAACTGAGC	2628
Db	2521	TTTATTTTAAATAGGTTTGTCGCTTAATCTTCAATAGAATTTAAATATGTCACAACTGAGC	2580
Qy	2629	ACAAATTAATTAATTTAATTAATTAACAAATAGC	2661
Db	2581	ACAAATTAATTAATTTAATTTAATTAACAAATAGC	2613

RESULT 3	BD156534	2639 bp	DNA	linear	PAT 17-JAN-2003
LOCUS	BD156534				
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD156534				
VERSION	BD156534.1	GI:27862292			
KEYWORDS	JP 2002191363-A/11377.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE 1 (bases 1 to 2639)  
AUTHORS Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J...  
TITL Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
JOURN Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 11977 09-JUN-2002;

COMMENT	OS	HELIX RESEARCH INSTITUTE
	Homo sapiens (human)	

PN JP 2002191363-A/11377  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
PI SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
(79) . . (2034) .  
FT CDS

FEATURES	source	Location/Qualifiers
1..2639		
/organism="Homo sapiens"		
/mol_type="genomic DNA"		
/db_xref="taxon:9606"		

	Query Match	Score	DB	Length
Best Local Similarity	79.6%	2133	6	2639
Matches	2133	Conservative	0	Mismatches 0; Indels 0; Gaps 0;

QY	649	CAACCTCGGCGCTCTTCCACAGGACGCTTCTCTCAAGGCGAGAGCGCTCTCAAGTTGT	708
Db	601	ACAGCTTCGGCGCTTTTCCAGGAGCGCTTCTCTCAAGGCGAGAGCGCTCTCAAGTTGT	660
QY	709	GGGCTCTGGGGGTGGTGGCATCTCCCTTAAGAAAGTGGTCCCTCTGTGGCAAGTGATGA	768
Db	661	GGGTCTGGGGGTGGTGGCATCTCCCTTAAGAAAGTGGTCCCTCTGTGGCAAGTGATGA	720
QY	769	GTCTTCAGCTTTTGCTCTAGCTCTCCCAAGACAGAAAGACTGGTCTTTTGTCTGTGGAA	828
Db	721	GTCTTCAGCTTTTGCTCTAGCTCTCCCAAGACAGAAAGACTGGTCTTTTGTCTGTGGAA	780
QY	829	CCCCAGGATCAAGAGAAATTGGAGCCCGTGAAGAAATAATGAGAGGAGATGGGAGACTT	888
Db	781	CCCCAGGATCAAGAGAAATTGGAGCCCGTGAAGAAATAATGAGAGGAGATGGGAGACTT	840
QY	889	GACCTGAACGGGCGAGTTGTTGGTGGCACAACCGCGTAGAATGCCCAACCGTCCAAGAG	948
Db	841	GACCTGAACGGGCGAGTTGTTGGTGGCACAACCGCGTAGAATGCCCAACCGTCCAAGAG	900
QY	949	GACCTCAAGAGCAGCGGCTGGAGACCAAGATGAGAGAGCTGACATGCAATATC	1008
Db	901	GACCTCAAGAGCAGCGGCTGGAGACCAAGATGAGAGAGAGCTGACATGCAATATC	960
QY	1009	TGGCAGAGACCTGCTGCACAGACTGGGTAGTTTGAAGCCCTGCATAGCAACGTTTGGGCG	1068
Db	961	TGGCAGAGACCTGCTGCACAGACTGGGTAGTTTGAAGCCCTGCATAGCAACGTTTGGGCG	1020
QY	1069	GCTTGCTACTCGGGCTGGATGGAACGCGTCGTCCTGTGTCTTACCTGCGCGTGCCTGTG	1128
Db	1021	GCTTGCTACTCGGGCTGGATGGAACGCGTCGTCCTGTGTCTTACCTGCGCGTGCCTGTG	1080
QY	1129	GAGCGGATCTGTAAAAACAATCTCAACAACCTGTGAAAGATACCTTATCAAGCAT	1188
Db	1081	GAGCGGATCTGTAAAAACAATCTCAACAACCTGTGAAAGATACCTTATCAAGCAT	1140
QY	1189	CCAGACAAAGATCGCAGTGAAGAAAGTGTGAAGTATGATGCCAAGAAATAAATCACT	1248
Db	1141	CCAGACAAAGATCGCAGTGAAGAAAGTGTGAAGTATGATGCCAAGAAATAAATCACT	1200
QY	1249	CAAGACATGCTGCAGACCCAAAGTACAGGCGGCTTTTCTGATGAAGAAAGGAGTTCAAG	1308
Db	1201	CAAGACATGCTGCAGACCCAAAGTACAGGCGGCTTTTCTGATGAAGAAAGGAGTTCAAG	1260
QY	1309	GACCTGTGAGCTGTCAAGCGTTGACAGTAGAGTCTTCAACATTAAGCCAGCAATACGTC	1368
Db	1261	GACCTGTGAGCTGTCAAGCGTTGACAGTAGAGTCTTCAACATTAAGCCAGCAATACGTC	1320
QY	1369	GTGTGCCGGCATGTCTCTAGATCAAGAAAGCAGCGGCGCAGCGCTCCCATCTGCCAGCA	1428
Db	1321	GTGTGCCGGCATGTCTCTAGATCAAGAAAGCAGCGGCGCAGCGCTCCCATCTGCCAGCA	1380
QY	1429	CCCCAGGCGGACAGAGAGCCCAACAGGCCCTGGGGGATGACACCTCACAGTCCGTAGC	1488
Db	1381	CCCCAGGCGGACAGAGAGCCCAACAGGCCCTGGGGGATGACACCTCACAGTCCGTAGC	1440
QY	1489	CTGAACAGCAGCTCAAGATTACGTGTGCCCTGTGCAGAGAACCAAGCCCTGTGCACC	1548
Db	1441	CTGAACAGCAGCTCAAGATTACGTGTGCCCTGTGCAGAGAACCAAGCCCTGTGCACC	1500
QY	1549	TGCTGCTTCAAGCCCATATGCCAGCCGAGAGAGCGAGCAGAGCAACCGCGTGTGCC	1608
Db	1501	TGCTGCTTCAAGCCCATATGCCAGCGAGAGAGCGAGCAGAGCAACCGCGTGTGCC	1560
QY	1609	CCTCAGCAGTGTGGGTGTGCTGCTGCACAGCTTTCTGCAACCTGTATCTGGGGGTGACCCGG	1668
Db	1561	CCTCAGCAGTGTGGGTGTGCTGCTGCACAGCTTTCTGTATCTGGGGGTGACCCGG	1620
QY	1669	ACCGGCTCTACGGCTGCTGGCCCCCGTTTGTAGCTCAACCTGGGTGACAAATGTCTG	1728
Db	1621	ACCGGCTCTACGGCTGCTGGCCCCCGTTTGTATAGCTCAACCTGGGTGACAAATGTCTG	1680
QY	1729	GACGCGGTCTGAACAACAAGCTTACAGAGTCAACATCTTGAAGAAATTAATCTGGCAAC	1788

```

Db      1681 GAGCGGCTGTGACACACACAGCTACAGTCAAGACATCTTGAAGATTACCTGGCAACC 1740
      |||
Qy      1789 AGAGGTTTGCATGGAAGAAACATGTTGACCGAGAGCGCTCGTGGCTCTCCAGCGGGAGTG 1848
      |||
Db      1741 AGAGGTTTGCATGGAAGAAACATGTTGACCGAGAGCGCTCGTGGCTCTCCAGCGGGAGTG 1800
      |||
Qy      1849 TTTTCTGCTGTCTGATTACAGAGTCAAGGAGACACCGTTCTGTGTTTACTGCTGTGGCTTG 1908
      |||
Db      1801 TTTTCTGCTGTCTGATTACAGAGTCAAGGAGACACCGTTCTGTGTTTACTGCTGTGGCTTG 1860
      |||
Qy      1909 CCGAGCTTCCGAGAGTGAAGCTTACAGTATCCGAGACAAATCTCTGTTCCGAGTTGGCA 1968
      |||
Db      1861 CCGAGCTTCCGAGAGTGAAGCTTACAGTATCCGAGACAAATCTCTGTTCCGAGTTGGCA 1920
      |||
Qy      1969 GTGGCCGTAACATCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028
      |||
Db      1921 GTGGCCGTAACATCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
      |||
Qy      2029 GCTTCCACAGCCCATGGAATTTCAATCTATCTGTGAACACACAGGTTCAAAAACCTAGCA 2088
      |||
Db      1981 GCTTCCACAGCCCATGGAATTTCAATCTATCTGTGAACACACAGGTTCAAAAACCTAGCA 2040
      |||
Qy      2089 TCCAGAGGCGCTGAGAGCGCTTTCAGACCTGAGAGTGAAGAGCGCTGTTTAAATATACA 2148
      |||
Db      2041 TCCAGAGGCGCTGAGAGCGCTTTCAGACCTGAGAGTGAAGAGCGCTGTTTAAATATACA 2100
      |||
Qy      2149 GAGACACAGACGTCAGAGTGTGTTTTCAGAGCCCTGAGAGGAAAGAGACGAGGCTCTCCGA 2208
      |||
Db      2101 GAGACACAGACGTCAGAGTGTGTTTTCAGAGCCCTGAGAGGAAAGAGACGAGGCTCTCCGA 2160
      |||
Qy      2209 CAGGCTCTGAGGAGTGAAGCTTCTGTGAGAGCTTTTAAACCTGAGTGAAGCCCTCCCA 2268
      |||
Db      2161 CAGGCTCTGAGGAGTGAAGCTTCTGTGAGAGCTTTTAAACCTGAGTGAAGCCCTCCCA 2220
      |||
Qy      2269 GAGCCCGGAGGAGCGAGCGCCCTCTGTGAGAGCGCTGAGGAGCGCTGAGTGAAGCT 2328
      |||
Db      2221 GAGCCCGGAGGAGCGAGCGCCCTCTGTGAGAGCGCTGAGGAGCGCTGAGTGAAGCT 2280
      |||
Qy      2329 CAGCAGCAGAGACAGAGCCCTTCTGTGAACATGCGCGCTCCCGCCAGAGGAGGAGCTTT 2388
      |||
Db      2281 CAGCAGCAGAGACAGAGCCCTTCTGTGAACATGCGCGCTCCCGCCAGAGGAGGAGCTTT 2340
      |||
Qy      2389 GCTCTTTTGTACATTTCCGAAACCTACAGTAAAGAGAGCTGTTTTCAGAGAAAGTT 2448
      |||
Db      2341 GCTCTTTTGTACATTTCCGAAACCTACAGTAAAGAGAGCTGTTTTCAGAGAAAGTT 2400
      |||
Qy      2449 TCAAGGAGAGAGGAGCAAGTTTATCAAAAACATTTGTTTTCAGAGAAAGGAGCATTAAGTTA 2508
      |||
Db      2401 TCAAGGAGAGAGGAGCAAGTTTATCAAAAACATTTGTTTTCAGAGAAAGGAGCATTAAGTTA 2460
      |||
Qy      2509 CAGCCTACAGAGACCTACACAAATCTCTGCTGCGGAGAAACACACAGCATTTTATCTAATT 2568
      |||
Db      2461 CAGCCTACAGAGACCTACACAAATCTCTGCTGCGGAGAAACACACAGCATTTTATCTAATT 2520
      |||
Qy      2569 TTTTATTTATAGTTGTTGCTGTTATCTCTATATAGATTTTAAATGTCAAACTGTAGC 2628
      |||
Db      2521 TTTTATTTATAGTTGTTGCTGTTATCTCTATATAGATTTTAAATGTCAAACTGTAGC 2580
      |||
Qy      2629 ACAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2661
      |||
Db      2581 ACAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2613
      |||

```

```

RESULT 4
AK027687
LOCUS      AK027687      2639 bp      mRNA      linear      PRI 30-JAN-2004
DEFINITION Homo sapiens CDNA FLJ14781 f1s, clone NT2RPA000455, weakly similar
              to TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0.
ACCESSION AK027687.1 GI:14042552
VERSION   AK027687.1
KEYWORDS  oligo capping; f1s (full insert sequence).
SOURCE    Homo sapiens (human)

```

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakematsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Oobayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikawa,E., Omura,Y., Abe,K., Kamihara,K., Katsura,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Itoh,H., Kimata,M., Watanabe,M.,
Murakawa,K., Fujimori,K., Tanai,H., Kimura,M., Watanabe,M.,
Hirooka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Hoshino,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiyasu,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Mueshino,K., Yuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohara,N., Sano,S., Moriya,S., Momiyama,H., Satch,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kunagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujitara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigetake,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togaishi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuno,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE
AUTHORS
2
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,
Matsunawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakematsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Sasaki,N.
NEBO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE
AUTHORS
3 (bases 1 to 2639)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction,
5' & 3'-end one pass sequencing and clone selection: Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Source
1..2639
Location/Qualifiers
1..2639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RPA000455"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RPA"
/note="Cloning vector: pME18SF3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
79..2037
/note="unnamed protein product"
/codon_start=1
/protein_id="BA55297.1"

```

CDS





Db 531 AAGAGCTGGAGCTCTTCCAGAGAGCTTCTCGGAGGAGAGAGGCTTCTCAGATTGT 590  
Qy 709 GGGTCTGGGGGTGGTGGCAATCTCCCTAAAGAAAGTGGTCCCTCTGTGGCAAGTATGAA 768  
Db 591 GGGTCTGGGGGTGGTGGCAATCTCCCTAAAGAAAGTGGTCCCTCTGTGGCAAGTATGAA 650  
Qy 769 GTCTCAGGCTTTGGCTCAGCTCTCCGAGACAGAAAGAGTGGCTCTTTTGTGTGGAA 828  
Db 651 GTCTCAGGCTTTGGCTCAGCTCTCCGAGACAGAAAGAGTGGCTCTTTTGTGTGGAA 710  
Qy 829 CCCAGAGATCAGAGAGATTGGAGCCCGTGAAGAAATAGAGAGAGATGGGAGCTT 888  
Db 711 CCCAGAGATCAGAGAGATTGGAGCCCGTGAAGAAATAGAGAGAGATGGGAGCTT 770  
Qy 889 GACCTGAACGGGAGATTGTTGGTGGCAACCCCGTGAAGAAATGCCAAACCTGTCCAGAG 948  
Db 771 GACCTGAACGGGAGATTGTTGGTGGCAACCCCGTGAAGAAATGCCAAACCTGTCCAGAG 830  
Qy 949 GACCTGAACGGGAGATTGTTGGTGGCAACCAAGATGAGAGAGAGCTGAATGCAATCATC 1008  
Db 831 GACCTGAACGGGAGATTGTTGGTGGCAACCAAGATGAGAGAGAGCTGAATGCAATCATC 890  
Qy 1009 TGCAGAGACCTGCTGCAACGATGCTGAGATTGGCAGCCCTGCAAGCAAGTTCGCGG 1068  
Db 891 TGCAGAGACCTGCTGCAACGATGCTGAGATTGGCAGCCCTGCAAGCAAGTTCGCGG 950  
Qy 1069 GCTTGTCTACTCGGGCTGAGATGAGAGCGCTGTCTGTGTCTACTCGGCGCTGTCCG 1128  
Db 951 GCTTGTCTACTCGGGCTGAGATGAGAGCGCTGTCTGTGTCTACTCGGCGCTGTCCG 1010  
Qy 1129 GAGCGGATCTGTAAACCAATCCTCAACACTGTGTGAAGCAATCCTCAATCCAGCAT 1188  
Db 1011 GAGCGGATCTGTAAACCAATCCTCAACACTGTGTGAAGCAATCCTCAATCCAGCAT 1070  
Qy 1189 CCAGACAAAGATCGCAGTGAAGAAAGTGTGCAAAAGTATGATGCCAGATTAATCACT 1248  
Db 1071 CCAGACAAAGATCGCAGTGAAGAAAGTGTGCAAAAGTATGATGCCAGATTAATCACT 1130  
Qy 1249 CAAGACATGCTGCAAGCCCAAGTCAAGCGGCTCTTTTCTGATGAAGAGAGATTCAAG 1308  
Db 1131 CAAGACATGCTGCAAGCCCAAGTCAAGCGGCTCTTTTCTGATGAAGAGAGATTCAAG 1190  
Qy 1309 GACCTGTGAGAGCTGCAAGCTTGAAGATGCTCTGAGCAATTGACAGGCATACGTC 1368  
Db 1191 GACCTGTGAGAGCTGCAAGCTTGAAGATGCTCTGAGCAATTGACAGGCATACGTC 1250  
Qy 1369 GTGTGCGGCAAGTCTCTGATGACAAAGAGAGCGGCGCAGCTCCCACTGCCAGCA 1428  
Db 1251 GTGTGCGGCAAGTCTCTGATGACAAAGAGAGCGGCGCAGCTCCCACTGCCAGCA 1310  
Qy 1429 CCCGAGGCGAGAGCCAGAGAGCCCAAGAGCCCTGTGGGAGATGCACTCTCAAGTCCGTCAGC 1488  
Db 1311 CCCGAGGCGAGAGCCAGAGAGCCCAAGAGCCCTGTGGGAGATGCACTCTCAAGTCCGTCAGC 1370  
Qy 1489 CTGACGACAGCAGTCCAGATTAACGTGTGCTCTGTGCAAGAAAGCAAGCCCTGTGACAC 1548  
Db 1371 CTGACGACAGCAGTCCAGATTAACGTGTGCTCTGTGCAAGAAAGCAAGCCCTGTGACAC 1430  
Qy 1549 TGTGTCTTCAAGCCCAATGCCCCAGAGAGCGGAGAGCGAGAGAGAGCCCGCGTGTGCGC 1608  
Db 1431 TGTGTCTTCAAGCCCAATGCCCCAGAGAGCGGAGAGCGGAGAGAGAGCCCGCGTGTGCGC 1490  
Qy 1609 CCTCAGACAGTGTGGGTGTGCTGTGAGCCCTTCTGCACTGTACTGTGGAGCTGACCCG 1668  
Db 1491 CCTCAGACAGTGTGGGTGTGCTGTGAGCCCTTCTGCACTGTACTGTGGAGCTGACCCG 1550  
Qy 1669 ACCGCTGTGACAGCGCTGCGCCGCTTTTGTGAAGTCAACCTGTGGTGAACAAGTCTG 1728  
Db 1551 ACCGCTGTGACAGCGCTGCGCCGCTTTTGTGAAGTCAACCTGTGGTGAACAAGTCTG 1610  
Qy 1729 GACGCGTGTGACAAACAAGCTACAGATCAGATCCTGAAAGAAATTAACCTGCAAC 1788

Db 1611 GACGCGTGTGACAAACAAGCTACAGATGACATCTGAAAGAAATTAACCTGCAAC 1670  
Qy 1789 AAGAGTTTGAATGAGAAACAATGTGACCGAGAGCTGTGGCTCTCCAGGGGAGAG 1848  
Db 1671 AAGAGTTTGAATGAGAAACAATGTGACCGAGAGCTGTGGCTCTCCAGGGGAGAG 1730  
Qy 1849 TTTCTGTCTGTGATTAAGAGTCAAGAGAGACCGTCTGTGTTACTGTGTGGCTG 1908  
Db 1731 TTTCTGTCTGTGATTAAGAGTCAAGAGAGACCGTCTGTGTTACTGTGTGGCTG 1790  
Qy 1909 GCGAGCTTCCGTGAGCTGACCTATGATGTGGAGAAACATTTCTGTCCGAGTTGCCA 1968  
Db 1791 GCGAGCTTCCGTGAGCTGACCTATGATGTGGAGAAACATTTCTGTCCGAGTTGCCA 1850  
Qy 1969 GTGGCCGTGAATCCCGTCTGACCTGCTATCTGGGCGCGTAATGCGCATCAGGTGAA 2028  
Db 1851 GTGGCCGTGAATCCCGTCTGACCTGCTATCTGGGCGCGTAATGCGCATCAGGTGAA 1910  
Qy 2029 GCTCAACAGCCATGAATTAATCAATATCTGTGAACAGACAAGGTTCAAAAATAAGA 2088  
Db 1911 GCTCAACAGCCATGAATTAATCAATATCTGTGAACAGACAAGGTTCAAAAATAAGA 1970  
Qy 2089 TCCAGAGCCCTGAGCAGCTTTCAAGCATGAGGTGAAGAGAGCGTGTTTAAATACA 2148  
Db 1971 TCCAGAGCCCTGAGCAGCTTTCAAGCATGAGGTGAAGAGAGCGTGTTTAAATACA 2030  
Qy 2149 GAGACAAAGCAGTCAAGTGTGTTTCAAGCCCCCTGAAGGAGAGAGAGGTCTCCCA 2208  
Db 2031 GAGACAAAGCAGTCAAGTGTGTTTCAAGCCCCCTGAAGGAGAGAGAGGTCTCCCA 2090  
Qy 2209 CAGGTGCTGTGGGAGTACTCTGTGTGAGCTTTTAAACCTCTGAGTGAAGCCCTCCCA 2268  
Db 2091 CAGGTGCTGTGGGAGTACTCTGTGTGAGCTTTTAAACCTCTGAGTGAAGCCCTCCCA 2150  
Qy 2269 GAGCCCGGAGGCGCAGAGCCGCTCTGTGTGAGAGCTGTGGCAGAGGCTGTGTGGCAT 2328  
Db 2151 GAGCCCGGAGGCGCAGAGCCGCTCTGTGTGAGAGCTGTGTGGCAGAGGCTGTGTGGCAT 2210  
Qy 2329 CAGCAGCAGAGCAGAGCTTTCTGTGAACATGCGGCGCTGCGGAGAGAGGAGGCACTTT 2388  
Db 2211 CAGCAGCAGAGCAGAGCTTTCTGTGAACATGCGGCGCTGCGGAGAGAGGAGGCACTTT 2270  
Qy 2389 GCTCTTTGTATCAATTTCCGAAACTACAGTTAAAGAGAGAGTCTGTTTCAAGAAAGTT 2448  
Db 2271 GCTCTTTGTATCAATTTCCGAAACTACAGTTAAAGAGAGAGTCTGTTTCAAGAAAGTT 2330  
Qy 2449 TCAAGGAGAGAGGAGAGTATTAATCAAAAATTTGTTTCAAGAGAGAGCAATTAATT 2508  
Db 2331 TCAAGGAGAGAGGAGAGTATTAATCAAAAATTTGTTTCAAGAGAGAGCAATTAATT 2390  
Qy 2509 CAGCCTACAGAGCAGTACACATATCTGTGCTGTGGAGAAACAGAGCATTTTATCTATT 2568  
Db 2391 CAGCCTACAGAGCAGTACACATATCTGTGCTGTGGAGAAACAGAGCATTTTATCTATT 2450  
Qy 2569 TTTATTTTAAATAGTGTGTGCTTATCTTAAATTAAGATTTAAATGTCACAACTGTAGC 2628  
Db 2451 TTTATTTTAAATAGTGTGTGCTTATCTTAAATTAAGATTTAAATGTCACAACTGTAGC 2510  
Qy 2629 ACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2661  
Db 2511 ACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2543

RESULT 6  
AX405642  
LOCUS AX405642 3181 bp DNA linear PAT 14-JUN-2002  
DEFINITION Sequence 57 from Patent W00222660.  
ACCESSION AX405642  
VERSION AX405642.1 GI:21438723  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



REFERENCE  
1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
AUTHORS  
Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,  
Xue, A.J., Yang, Y., Wehrman, T., and Drmanac, R.T.  
TITLE  
Novel nucleic acids and polypeptides  
JOURNAL  
Patent: WO 022660-A 57 21-MAR-2002;  
HYSEQ, INC. (US)  
FEATURES  
source  
1. .3181  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
96..1967  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD34749.1"  
/db\_xref="GI:21438724"  
/translation="MERPEKQSPPOPMGRLLRLGAREGSPHYLLRKEWTIGRR  
GCDLSPSKLYSGDHCRIVYDEKSGQVLEDTSTSGTYINKLKYKQTCPLQTDV  
IIVYRKNEPEHNVATLYESLSBKQMTSEFEMWPCVAQGLKLGSSDPTLASQ  
SIVITSGGGGGLSPKSGPSVASDEVSPASALPDKTASFSLSLEPQDQEDLEPVKK  
MEGDDLDLNGOLIVAPPRNAQVHEIDVRAAGKPMKMETLTCIIICODLHDCVSL  
OPEMTFCAACYSMMERSLCPCTCPYERLCXNHLINLVAVYLIOHPDSEED  
VOSMARNKITQDMLQKYRSPDESGSSEDLRLSDSRSNDISQPYVRCQPE  
YRROAAPHCHAPBEGAPALGDAPPTSVITAVDYVCPQGSALCTCCFP  
MDRAERQDPRVAPQCCAVLCOPCHLFWGCTRGCTGCLAPCELMGDKCDGV  
LNNSTYEDILNKYLAIRGLTWKNMLTESLVALQRFVALSLQYRATGDIIVLCYCGLR  
SRELIYQYQONIPASELPVATSRDPCWGNCRCTQVVAHAKMFNHCIEQTRKN"

ORIGIN

Query Match	Best Local Similarity	Score	Pred.	No. 0;	Mismatches	Indels	Gaps
709	GGGTCGGGGGCTGGTCATCTCCCTAAAGAAAGTGTCCCTGTGGCAAGTGA	60.4%;	99.7%;	1618;	DB 6;	Length 3181;	
591	GGGTCGGGGGCTGGTCATCTCCCTAAAGAAAGTGTCCCTGTGGCAAGTGA						
769	GTCTCGAGCTTTGGCTCAGCTCTCCCAAGCAAGAAAGTGTCTTTTGTGGTGA						
651	GTCTCGAGCTTTGGCTCAGCTCTCCCAAGCAAGAAAGTGTCTTTTGTGGTGA						
829	CCCGAGATCAGAGAGATTGGAGCCCGGAGAAAGAAATGAGAGAGAGAGAGAGT						
711	CCCGAGATCAGAGAGATTGGAGCCCGGAGAAAGAAATGAGAGAGAGAGAGAGT						
889	GACTGGAAGGAGAGTGTGGTGGCAACCGCTGAGAAATGCCCAACCGTCCACGAG						
771	GACTGGAAGGAGAGTGTGGTGGCAACCGCTGAGAAATGCCCAACCGTCCACGAG						
949	GACGTCAAGAGAGAGAGTGTGGAGCCAGCAAGATGAGAGAGAGAGAGAGAGT						
831	GACGTCAAGAGAGAGAGTGTGGAGCCAGCAAGATGAGAGAGAGAGAGAGAGT						
1009	TCCAGAGAGAGTGTGGAGAGTGTGGAGAGTGTGGAGAGAGAGAGAGAGAGT						
891	TCCAGAGAGAGTGTGGAGAGTGTGGAGAGTGTGGAGAGAGAGAGAGAGAGT						
1069	GCTTGTCTACTCGAGCTGAGTGGAGCGCTGTCTGTGTCTTACCTGTGCGCTGCT						
951	GCTTGTCTACTCGAGCTGAGTGGAGCGCTGTCTGTGTCTTACCTGTGCGCTGCT						
1129	GAGCGGATTTTAAAAACCAATCTCTCAACAACCTGTGGAGAGATCTTCAACGAT						
1011	GAGCGGATTTTAAAAACCAATCTCTCAACAACCTGTGGAGAGATCTTCAACGAT						
1189	CCAGACAGAGTGTGGAGAGAGTGTGGAGAGAGTGTGGAGAGAGAGAGAGAGT						
1071	CCAGACAGAGTGTGGAGAGAGTGTGGAGAGAGTGTGGAGAGAGAGAGAGAGT						
1249	CAAGACATGTCTGAGCCCAAGTGTGGAGCGCTTTTGTGTGATGAGAGAGAGT						
1131	CAAGACATGTCTGAGCCCAAGTGTGGAGCGCTTTTGTGTGATGAGAGAGAGT						

QY	1309	GACTGCTGAGAGCTGTGACAGCTTGTGACAGTGTCTTGTGACATTTAGCCAGCATACGTC	1368
DB	1191	GACTGCTGAGAGCTGTGACAGCTTGTGACAGTGTCTTGTGACATTTAGCCAGCATACGTC	1250
QY	1369	GTGTGCTGGAGAGTGTCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1428
DB	1251	GTGTGCTGGAGAGTGTCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1310
QY	1429	CCGAGT	1488
DB	1311	CCGAGT	1370
QY	1489	CTGACAGAGAGAGTGTGACAGATTTACGTGTGCTTGTGACAGAGAGAGAGAGT	1548
DB	1371	CTGACAGAGAGAGTGTGACAGATTTACGTGTGCTTGTGACAGAGAGAGAGAGT	1430
QY	1549	TGCTGCTTCCAGT	1608
DB	1431	TGCTGCTTCCAGT	1490
QY	1609	CCTGACAGAGTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1668
DB	1491	CCTGACAGAGTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1550
QY	1669	ACCGGCTGCTACGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1728
DB	1551	ACCGGCTGCTACGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1610
QY	1729	GACGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1788
DB	1611	GACGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1670
QY	1789	AGAGGTTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1848
DB	1671	AGAGGTTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1730
QY	1849	TTTCTGCTGTCTGATTTACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGT	1908
DB	1731	TTTCTGCTGTCTGATTTACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGT	1790
QY	1909	GCGAGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1968
DB	1791	GCGAGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1850
QY	1969	GTGGCCGTGATCCTCGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	2028
DB	1851	GTGGCCGTGATCCTCGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT	1910
QY	2029	GCTCAGCAGT	2088
DB	1911	GCTCAGCAGT	1970
QY	2089	TCCAGT	2148
DB	1971	TCCAGT	2030
QY	2149	GAGACAGT	2208
DB	2031	GAGACAGT	2090
QY	2209	CAGGTGCTGTGGAGTGTCTTGTGTGAGAG--TTTACCTCTGTGTGTGTGTGT	2248
DB	2091	CAGGTGCTGTGGAGTGTCTTGTGTGAGAG--TTTACCTCTGTGTGTGTGTGT	2148
QY	2269	GAGCCCGAGT	2328
DB	2149	GAGCCCGAGT	2208
QY	2329	CAGCAGT	2388
DB	2209	CAGCAGT	2268





```

Db      1760 CGAGGCTTCGTAAGCTGACCTAATCATATCAGACAGAACTCTCTCCGAGTCCCA 1819
QY      1969 GTGGCCGTAAATCATCCGCTCTGACTGCTAGCGGCGGTAACTGCCGCACTCAGGTGAA 2028
Db      1820 GTGGCCGTAAATCATCCGCTCTGACTGCTAGCGGCGGTAACTGCCGCACTCAGGTGAA 1879
QY      2029 GCTTACCAAGCCATGAAATTCATATCTGTTGAAACAGACAGGTTCAAAAATAAGCA 2088
Db      1880 GCTTACCAAGCCATGAAATTCATATCTGTTGAAACAGACAGGTTCAAAAATAAGCA 1939
QY      2089 TCCAGAGGCCCTGAGAGCTTTGAGACTGAGAGTGAAGAGAGGTTTAAATATCA 2148
Db      1940 TCCAGAGGCCCTGAGAGCTTTGAGACTGAGAGTGAAGAGAGGTTTAAATATCA 1999
QY      2149 GAGACAGACAGTCAAGGTGTTTCAAGCCCTGAGAGGAAAGGAGCGAGGCTCTCCA 2208
Db      2000 GAGGCAAGACAGTCAAGGTGTTTCAAGCCCTGAGAGGAAAGGAGCGAGGCTCTCCA 2059
QY      2209 CAGGTGCTCTGAGGAGTACTTCTGTGAGAGCTTTTAACTCTGAGTGAAGCTTCCCA 2268
Db      2060 CAGGTGCTCTGAGGAGTACTTCTGTGAGAGCTTTTAACTCTGAGTGAAGCTTCCCA 2117
QY      2269 GAGCCCGGGGGCCGAGCCGCTCTGTGAGAGGCTGAGGAGGAGTCTGTGAGTGCAT 2328
Db      2118 GAGCCCGGGGGCCGAGCCGCTCTGTGAGAGGCTGAGGAGGAGTCTGTGAGTGCAT 2177
QY      2329 CAGCAGCAGACAGAAAGCTTTCTGTGAAACAGCGCCGCTCCGAGAGGAGGAGTCTT 2388
Db      2178 CAGCAGCAGACAGAAAGCTTTCTGTGAAACAGCGCCGCTCCGAGAGGAGGAGTCTT 2237
QY      2389 GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGAGAGTCTGTTTCAAGAAAGTT 2448
Db      2238 GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGAGAGTCTGTTTCAAGAAAGTT 2297
QY      2449 TCAGAGGAGAGGAGCAAGTTTCAAAAATGTTTCAAGAGAGAGAGCAATAGTTTA 2508
Db      2298 TCAGAGGAGAGGAGCAAGTTTCAAAAATGTTTCAAGAGAGAGAGCAATAGTTTA 2357
QY      2509 CAGCCTACAGACGTAACATATCTCTGCTGAGGAAAAACAAGCATTTTATCTATT 2568
Db      2358 CAGCCTACAGACGTAACATATCTCTGCTGAGGAAAAACAAGCATTTTATCTATT 2417
QY      2569 TTTATTTATATGTTTGTGCTTATCTCTTAAATTAAGTTAAAGTCAACAAGTGC 2628
Db      2418 TTTATTTATATGTTTGTGCTTATCTCTTAAATTAAGTTAAAGTCAACAAGTGC 2477
QY      2629 ACAATATATATATTTATATTTTACAAATGAC 2661
Db      2478 ACAATATATATATTTATATTTTACAAATGAC 2510

RESULT 8
BD156548      3138 bp      DNA      linear      PAR 17-JUN-2003
LOCUS      BD156548
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD156548
VERSION      BD156548.1 GI:27862306
KEYWORDS      JP 2002191363-A/11391.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3138)
Ota,T., Iwagai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K., and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 11391 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/11391
PD      09-JUL-2002
PF      28-JUL-2000 JP 2002080990

```

```

PI      TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI      SAITO,
PI      JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI      KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC      C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(65) . . (1933) .
FT      CDS
source      1. 3138
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match      54.7%; Score 1465; DB 6; Length 3138;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
QY      709 GGGTCTGGGGGTGTGTGCGATCTCCCTTAAAGAAAGTGTCCCTGTGTGCAAGTATGA 768
Db      560 GGGTCTGGGGGTGTGTGCGATCTCCCTTAAAGAAAGTGTCCCTGTGTGCAAGTATGA 619
QY      769 GTTTCAGCTTTGCTCAGCTCTCCAGACAGAAAGACTGCGTCTTTTGTGTTGAA 828
Db      620 GTTTCAGCTTTGCTCAGCTCTCCAGACAGAAAGACTGCGTCTTTTGTGTTGAA 679
QY      829 CCCAGAGTCAAGAGAGATTTGAGAGCCCGTGAAGAAAGAAATGAGAGAGATGGGACCTT 888
Db      680 CCCAGAGTCAAGAGAGATTTGAGAGCCCGTGAAGAAAGAAATGAGAGAGATGGGACCTT 739
QY      889 GACCTGAACGGGCAAGTTTGTGTGTCAGCAACCGGTGAATGCCAAACGTCACAGAG 948
Db      740 GACCTGAACGGGCAAGTTTGTGTGTCAGCAACCGGTGAATGCCAAACGTCACAGAG 799
QY      949 GACGTCAAGACAGCGGCTGGGAAAGCAGACAAAGTGAAGAGAGCTGACATGATCATC 1008
Db      800 GACGTCAAGACAGCGGCTGGGAAAGCAGACAAAGTGAAGAGAGCTGACATGATCATC 859
QY      1009 TGGCAGAGCTGTGACAGACTGGGTGAGTTTGAAGCCCTGATGACACAGCTTCCGCGG 1068
Db      860 TGGCAGAGCTGTGACAGACTGGGTGAGTTTGAAGCCCTGATGACACAGCTTCCGCGG 919
QY      1069 GCTTGTCTACTCGGGGTGATGAGCGCTGTCTCTGTGTCTTACTGCGCTGTCCGCTG 1128
Db      920 GCTTGTCTACTCGGGGTGATGAGCGCTGTCTCTGTGTCTTACTGCGCTGTCCGCTG 979
QY      1129 GAGCGGATCTGTAAAAACCAATCTCTCAACAACTCTGTGAAGCATTAATCCAGCAT 1188
Db      980 GAGCGGATCTGTAAAAACCAATCTCTCAACAACTCTGTGAAGCATTAATCCAGCAT 1039
QY      1189 CCAAGCAAGATCTGAGAGAGAGATGTGAAAAGTATGATGACAGGAAATTAATCACT 1248
Db      1040 CCAAGCAAGATCTGAGAGAGAGATGTGAAAAGTATGATGACAGGAAATTAATCACT 1099
QY      1249 CAAAGCATGTCTGAGAGCCCAAGTCAAGCGGCTTTTCTGTGAAGAAAGGAGTTCAAG 1308
Db      1100 CAAAGCATGTCTGAGAGCCCAAGTCAAGCGGCTTTTCTGTGAAGAAAGGAGTTCAAG 1159
QY      1309 GACCTGTGAGAGCTGTCAAGAGTTGAAGATGCTTCAAGCATTAAGCCAGCCTATGCTC 1368
Db      1160 GACCTGTGAGAGCTGTCAAGAGTTGAAGATGCTTCAAGCATTAAGCCAGCCTATGCTC 1219
QY      1369 GTGTGCGGCGAGTGTCTGAGATACAGAAAGCAGGCGAGCTCCGCCCACTGCGCAGCA 1428
Db      1220 GTGTGCGGCGAGTGTCTGAGATACAGAAAGCAGGCGAGCTCCGCCCACTGCGCAGCA 1279
QY      1429 CCCAGAGGCGAGCCAGAGAGCCCAAGGCTTGGGGAGATGACCTCCACGCTCGTCAAGC 1488
Db      1280 CCCAGAGGCGAGCCAGAGAGCCCAAGGCTTGGGGAGATGACCTCCCAAGCTCGTCAAGC 1339

```



etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.

## FEATURES

Location/Qualifiers

1..3138

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="NT2RP4000648"

/cell\_line="NT2"

/cell\_type="teratocarcinoma"

/clone\_1ib="NT2RP4"

/note="Cloning vector: pME18SF13-mRNA from NT2 neuronal precursor cells after 2-weeks recinoic acid (RA) induction."

65..1936

## CDS

/note="unnamed protein product"

/codon\_start=1

/protein\_id="BA091817.1"

/db\_xref="GI:7023051"

/translation="MERPEEGKQSPPEQWGRLLRIGAEESGPHVLLRRRENTIGRR  
GCDLSPSNKLVSGHCRIVDEKSGQVLEDTSTGTVINKLKVKKQCTPLQTDV  
ILVYKNEPEHNVAYLVESEKQMTSEFEMVPCVAOAGKILASGSDPTLASQ  
SLVITSGGGGSLSPKSGPSVASDEVSPASALPDKTASFSLEPQDEDEPEKKK  
MEGDEDLINGOLIVAPRNQTVHEVRAAAGKPDMEETLTCIICODLADCVSL  
QCEMTFFCACTSGMERSLCTPCRCPEYERICKRHILNNVYATILIQPDKRSBED  
VQSMARKKITDMLQPKVRSFSDSESSDELLEISDVSSSDISQPYVRCQPE  
YRQQAQPHCPAPEGEPGAPOALGDPATSVSLTVAVDYVCPLOGSIALCTCFOP  
MYDRAREDDPRAVAPQCAVCIQPFCHLYMGCTGTGCAFPGLMLGKCDGV  
LNNNSYEDILNKYLATRGVATWKNMLTESVALORGVFLTSRYVADVLCCGGLR  
SFRELTYOYQONIPASELPVATSRDPVATSRDQVGNCRQYVAHHAMFNHICEGTRFKN"

## ORIGIN

Query Match 54.7%; Score 1455; DB 9; Length 3138;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1945; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY	709	GGGTCGGGGGTGGTGGCAATCTCCCTAAAGAAAGTGCCTCTGTGGCAAGTATGAA	768
DB	560	GGGTCGGGGGTGGTGGCAATCTCCCTAAAGAAAGTGCCTCTGTGGCAAGTATGAA	619
QY	769	GTCCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGATGCGTCTTTTGTGTTGAA	828
DB	620	GTCCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGATGCGTCTTTTGTGTTGAA	679
QY	829	CCCCAGATCAGAGAGATTTGAGCCCGTGAAGAAAGAAATGAGAGAGATGGGACCTT	888
DB	680	CCCCAGATCAGAGAGATTTGAGCCCGTGAAGAAAGAAATGAGAGAGATGGGACCTT	739
QY	889	GACCTGAAGGGGAGTTGTGTGCGCAACCGCGTGAAGAAAGTCCCAACCGTCCACGAG	948
DB	740	GACCTGAAGGGGAGTTGTGTGCGCAACCGCGTGAAGAAAGTCCCAACCGTCCACGAG	799
QY	949	GACCTCAGAGACGCGCTGGGAAAGCCAGACAAAGATGAGAGAGACGCTGACATGATCATC	1008
DB	800	GACCTCAGAGACGCGCTGGGAAAGCCAGACAAAGATGAGAGAGACGCTGACATGATCATC	859
QY	1009	TGCGAGGACCTGTGTGACGACGTCGCTGAGTTTGACGCTTGCATGACACGTTTTCGCGG	1068
DB	860	TGCGAGGACCTGTGTGACGACGTCGCTGAGTTTGACGCTTGCATGACACGTTTTCGCGG	919
QY	1069	GCTGTCTACTCGGGCTGATGAGAGCGTGTCCCTGTGTCTTACCTGCGCGCTGTCCCGG	1128
DB	920	GCTGTCTACTCGGGCTGATGAGAGCGTGTCCCTGTGTCTTACCTGCGCGCTGTCCCGG	979
QY	1129	GAGCGGATCTGTAAAAACCAATCTTCAACAACCTGTGGAAGCAATCTTCAATCCAGCAT	1188
DB	980	GAGCGGATCTGTAAAAACCAATCTTCAACAACCTGTGGAAGCAATCTTCAATCCAGCAT	1039
QY	1189	CCAGACAAAGATGCGAGTGAAGAAATGTGCAAAAGTATGATGCCAGAAATTAATCACT	1248
DB	1040	CCAGACAAAGATGCGAGTGAAGAAATGTGCAAAAGTATGATGCCAGAAATTAATCACT	1099
QY	1249	CAAGACATCTGACGCCCAAAAGTCAAGCGGCTCTTTTCTGATGAGAAAGAGTTCAAG	1308

DB	1100	CAAGACATCTGACACCCCAAAAGTCAAGCGGCTCTTTTCTGATGAAGAAAGGAGTTCAAG	1159
QY	1309	GACCTGTGAGCTGTGACAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1368
DB	1160	GACCTGTGAGCTGTGACAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	1219
QY	1369	GTCGCGGCAAGTGTCTGAGTGAAGAAAGGAGGCGGAGCTTCCCACTGCCAGCA	1428
DB	1220	GTCGCGGCAAGTGTCTGAGTGAAGAAAGGAGGCGGAGCTTCCCACTGCCAGCA	1279
QY	1429	CCCCGAGGCGAGCCAGAGAGCCCAAGAGCCCTTGGGGATGACCCCTCCAGTCCGTGAC	1488
DB	1280	CCCCGAGGCGAGCCAGAGAGCCCAAGAGCCCTTGGGGATGACCCCTCCAGTCCGTGAC	1339
QY	1489	CTGACGACAGAGTCTCAAGATTAAGTGTGCTCTGACAGGAAAGCAAGCTCTGTGAC	1548
DB	1340	CTGACGACAGAGTCTCAAGATTAAGTGTGCTCTGACAGGAAAGCAAGCTCTGTGAC	1399
QY	1549	TGCTGCTTCAAGCCCATGCGGACCGAGAGCGGAGCGGAGCAAGCTCTGTGAC	1608
DB	1400	TGCTGCTTCAAGCCCATGCGGACCGAGAGCGGAGCGGAGCAAGCTCTGTGAC	1459
QY	1609	CCTCAGAGTGTGCGGCTGTGCTGAGGCTTTTGTGACCTGTGACGAGGAGTGTGAC	1668
DB	1460	CCTCAGAGTGTGCGGCTGTGCTGAGGCTTTTGTGACCTGTGACGAGGAGTGTGAC	1519
QY	1669	ACCGGCTCTCAAGCTGCGGCTGCGGCTTTTGTGAGCTCAACTGTGAGTGAAGAGTGTG	1728
DB	1520	ACCGGCTCTCAAGCTGCGGCTGCGGCTTTTGTGAGCTCAACTGTGAGTGAAGAGTGTG	1579
QY	1729	GACGCGTGTGAAACAACAAGCTACAGTCAAGTCAATCTGAAAGATTAATCTGCAAC	1788
DB	1580	GACGCGTGTGAAACAACAAGCTACAGTCAAGTCAATCTGAAAGATTAATCTGCAAC	1639
QY	1789	AGAGGTTGATGAGAAACATGTTGACCGAGAGCCCTGCGTCTGACGAGGAGATG	1848
DB	1640	AGAGGTTGATGAGAAACATGTTGACCGAGAGCCCTGCGTCTGACGAGGAGATG	1699
QY	1849	TTTCTGCTGTGATTAAGAGTCAAGGAGACACCGTCTGTGTTACTGTGTTGAGCTG	1908
DB	1700	TTTCTGCTGTGATTAAGAGTCAAGGAGACACCGTCTGTGTTACTGTGTTGAGCTG	1759
QY	1909	GCGAGCTTCCGTGAGCTGACCTATGAGTATGAGGAAACATTTCTGCTTCCGAGTTGCCA	1968
DB	1760	GCGAGCTTCCGTGAGCTGACCTATGAGTATGAGGAAACATTTCTGCTTCCGAGTTGCCA	1819
QY	1969	GTCGCGGTAATCATCCCGCTGACGCTGACGAGGAGCGTAACTGCGGACCTCAGGTTGAA	2028
DB	1820	GTCGCGGTAATCATCCCGCTGACGCTGACGAGGAGCGTAACTGCGGACCTCAGGTTGAA	1879
QY	2029	GCTCAACAAGCATGAATTAATCATATCTGTGAACAGACAAAGGTTCAAAAATTAACA	2088
DB	1880	GCTCAACAAGCATGAATTAATCATATCTGTGAACAGACAAAGGTTCAAAAATTAACA	1939
QY	2089	TCCAGAGCCCTGTGAGCACTTTCAGCACTGAGAGTGAAGAGAGTGTGTTTAAATTAACA	2148
DB	1940	TCCAGAGCCCTGTGAGCACTTTCAGCACTGAGAGTGAAGAGAGTGTGTTTAAATTAACA	1999
QY	2149	GAGACAAACAGTCAAGGTTTTCACAGAGCCCTGAGGAGGAAAGGAGCGAGGTTCCGA	2208
DB	2000	GAGACAAACAGTCAAGGTTTTCACAGAGCCCTGAGGAGGAAAGGAGCGAGGTTCCGA	2059
QY	2209	CAGGTCTCTGAGGAGTCTTCTGTGAGGCTTTTACCTCTGAGTGAAGACCTTCCCA	2268
DB	2060	CAGGTCTCTGAGGAGTCTTCTGTGAGGCTTTTACCTCTGAGTGAAGACCTTCCCA	2117
QY	2269	GAGCCCGGAGGCGGAGCCGCTCTGAGTGAAGGCTGTGGGAGAGGCTCTGTGTGAGCAT	2328
DB	2118	GAGCCCGGAGGCGGAGCCGCTCTGAGTGAAGGCTGTGGGAGAGGCTCTGTGTGAGCAT	2177
QY	2329	CAGACAGAGACGAGGCTTTCTGTAACATGCGGCGCTCCCGCGGAGAGGAGGAGTTT	2388

Db	2178	CAGCAGCAGAGCGAAGCCTTCTGTAAACATCGCGCGCTCTGCGGAGAGGGGACGTTTT	2237
Qy	2389	GCTCTTTTGTGTCATTTTTCGAAACATACAGTTAAACACAAAGTCGTTTTTCAGGAAAAGTT	2448
Db	2238	GCTCTTTTGTACATTTTCCGAAACTACAGTTAAACGGAAAGTCGTTTTTCAGGAAAAGTT	2297
Qy	2449	TCAAGGAGAAAGGCGAAGTTTATCAAAAACATGTGTTTCAGGAGAAAGGAGCATAGTTTA	2508
Db	2298	TCAAGGAGAAAGGCGAAGTTTATCAAAAACATGTGTTTCAGGAGAAAGGAGCATAGTTTA	2357
Qy	2509	CAGCCTACAGACCGACACAAATATCTGTCGTGGGAAAACACAGCATTTTATCTATTT	2568
Db	2358	CAGCCTACAGACCGACACAAATATCTGTCGTGGGAAAACACAGCATTTTATCTATTT	2417
Qy	2569	TTTATTTTAAATAGTGTGTTATCTTCAATAGAATTTAAATGTCACAACTGAGC	2628
Db	2418	TTTATTTTAAATAGTGTGTCGTATCTTCAATAGAATTTAAATGTCACAACTGAGC	2477
Qy	2629	ACAAATATATATTTATTTAATTTTAAAAATGAC	2661
Db	2478	ACAAATATATATTTATTTAATTTTAAATGAC	2510

RESULT	10
AX834949	
LOCUS	AX834949
DEFINITION	Sequence 2073 from Patent EP1347046.
ACCESSION	AX834949
VERSION	AX834949.1 GI:35921084
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	2448 bp DNA
	linear
	PAT 15-DEC-2003

REFERENCE 1  
Iisoga, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Ngai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuho, Y.  
TITLE Full-length cDNA sequences  
Patent: EP 1347046-A 2013-24-SEP-2003;  
JOURNAL Research Association for Biotechnology (JP)  
FEATURES location/Qualifiers  
Source 1..2448

**ORIGIN**

	Query Match	35.2%	Score 944	DB 6	Length 2448	
	Best Local Similarity	99.9%	Pred. No. 0			
	Matches	994	Conservative	0	Mismatches	1
					Indels	0
					Gaps	0
Qy	708	TGGGCTCTGGGAGTGATGCATCTCCCTTAAAGAAAGTGCTCCCTGTGTGCAAGTATGA	767			
Db	357	TGGGCTCTGGGAGTGATGCATCTCCCTTAAAGAAAGTGCTCCCTGTGTGCAAGTATGA	416			
Qy	768	AGTCTCCAGCTTGTGCTCAGCTCTCTCCAGACAGAAAGACTGAGCTCTTGTGCTGTGA	827			
Db	417	AGTCTCCAGCTTGTGCTCAGCTCTCTCCAGACAGAAAGACTGAGCTCTTGTGCTGTGA	476			
Qy	828	ACCCAGAGATCAAGAGATTGTGAACCCGTGAAGAGAAATGAGAGAGATGAGGACCT	887			
Db	477	ACCCAGAGATCAAGAGATTGTGAACCCGTGAAGAGAAATGAGAGAGATGAGGACCT	536			
Qy	888	TGACCTGAACGGGACGTGTGTGTGTCACAAACCGGTGAATGCCCAACCGTCCACGA	947			
Db	537	TGACCTGAACGGGACGTGTGTGTGTCACAAACCGGTGAATGCCCAACCGTCCACGA	596			
Qy	948	GGAAGTCAAGAGCTGGGCTTGGGAAGCCAGACAAAGTGAAGAGAGACGTGACATGCATCAT	1007			
Db	597	GGAAGTCAAGAGCTGGGCTTGGGAAGCCAGACAAAGTGAAGAGAGACGTGACATGCATCAT	656			

QY	1008	CTGCAGAGACCTGTGTGACGACCTCGAGATTTCACACCTCGATGCAACGTTCTGCG	1067
Db	657	CTGCCAGGACCTGCTGCAACACATCGCGATGTTTGACGCCCTGCATGCAACAGTTCTGCGC	716
QY	1068	GCGTTTGCTACTCGGAGCTGAGATGAGCGCTCGTCCCTGTGTCTTACCTTGCGGCTGTCCGT	1127
Db	717	GCGTTGTGCTACTCGGAGCTGAGATGAGCGCTCGTCCCTGTGTCTTACCTTGCGGCTGTCCGT	776
QY	1128	GGAGCGGATCTGTATAAAACCAACCTCTCAACACTGTGTGAAGCAATCCTCATCCAGCA	1187
Db	777	GGAGCGGATCTGTATAAAACCAACCTCTCAACACTGTGTGAAGCAATCCTCATCCAGCA	836
QY	1188	TCCAGACAGAAGTGCAGTGAAGAAGATGTGCACAAAGTATGATGATGCCAGAAATAAAATCAC	1247
Db	837	TCCAGACAGAAGTGCAGTGAAGAAGATGTGCACAAAGTATGATGATGCCAGAAATAAAATCAC	896
QY	1248	TCAAGACATGTGTGACAGCCCAAAAGTCAGCGCGTCTTTTTCGTATGATGAAGAAGGAGTTTACA	1307
Db	897	TCAAGACATGTGTGACAGCCCAAAAGTCAGCGCGTCTTTTTCGTATGATGAAGAAGGAGTTTACA	956
QY	1308	GGAAGCTGTGAGGTGTGCAGACGTTTGAACGTGAAGTCTTCAGACATTAAGCCAGCATATAGT	1367
Db	957	GGAAGCTGTGAGGTGTGCAGACGTTTGAACGTGAAGTCTTCAGACATTAAGCCAGCATATAGT	1016
QY	1368	CGTGTGCCGGGCAGTGTCTCTGATGTCAGAAAGGACAGGCGGCGAGCTTCCCATCTGCCAGC	1427
Db	1017	CGTGTGCCGGGCAGTGTCTCTGATGTCAGAAAGGACAGGCGGCGAGCTTCCCATCTGCCAGC	1076
QY	1428	ACCGGAGGGGCGAGCCAGAGAGCCCAACAGGCGCTGTGGGGAATGACACCTTCACAGTCCGTCAG	1487
Db	1077	ACCGGAGGGGCGAGCCAGAGAGCCCAACAGGCGCTGTGGGGAATGACACCTTCACAGTCCGTCAG	1136
QY	1488	CCTGACGACAGACAGTCTCAGAGATTACGTGTGCGCTCTTGCACAGAAAGCCAGCCCTGTGCAC	1547
Db	1137	CCTGACGACAGACAGTCTCAGAGATTACGTGTGCGCTCTTGCACAGAAAGCCAGCCCTGTGCAC	1196
QY	1548	CTGCTGTCTTCACAGCCCATGCCCCGACCGGAGAGCGGAACGCGAGACGACACCCGCGTGTGCG	1607
Db	1197	CTGCTGTCTTCACAGCCCATGCCCCGACCGGAGAGCGGAACGCGAGACGACACCCGCGTGTGCG	1256
QY	1608	CCCTCAGACAGTGTGCGGATCTGCTGCAGAGCCTTTCTGCAACCTGTATCTGTGGGATCTGCACCG	1667
Db	1257	CCCTCAGACAGTGTGCGGATCTGCTGCAGAGCCTTTCTGCAACCTGTATCTGTGGGATCTGCACCG	1316
QY	1668	GACCGGCTGTACGCGCTGCTGCGGCCCGCTTTTGTG	1702
Db	1317	GACCGGCTGTACGCGCTGCTGCGGCCCGCTTTTGTG	1351

RESULT 11	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AK097671	2448 bp	mRNA	linear	PRI 30-JAN-2004					
AK097671		Homo sapiens cDNA FLJ30352 fis, clone TEST12033505, highly similar to Homo sapiens cell cycle checkpoint protein CHER mRNA.	AK097671						
AK097671.1	GI:21757518	oligo capping: fis (full insert sequence).							
		Homo sapiens							
		Homo sapiens (human)							
		Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
1		Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagaharui, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosohi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsura, N., Saco, K., Tanikawa, M., Yamazaki, M., Nishimura, Y., Ishibashi, T., Yamaehta, H., Murakawa, K., Fujimori, K., Tanai, H., Kinata, M., Watanabe, M., Hiraoaka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,							



Yosida, M., Horita, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsumura, H., Ichihara, T., Shibata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mitoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumejima, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabara, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Saitoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Maehiro, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuki, H., Hara, H., Sugiyama, T., Ito, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Negahori, K., Maehiro, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 2448)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genom@seihri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology (RAB) (supported by Japan Construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5- & 3-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

Location/Qualifiers

1..2448

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="TBS12033505"

/release\_type="releasable"

/clone\_id="TBS12"

/note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 35.2%; Score 944; DB 9; Length 2448;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 994; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 708 TGGGCTGGGGGTGGGATCTCCCTAAAGAAAGTGGTCCCTGTGGCAAGTGATGA 767

DB 357 TGGGCTGGGGGTGGGATCTCCCTAAAGAAAGTGGTCCCTGTGGCAAGTGATGA 416

QY 768 AGTCTCCAGCTTGGCTCAGCTCTCCCAACAGAAAGATGGTCTTTTCTGCTTGA 827

DB 417 AGTCTCCAGCTTGGCTCAGCTCTCCCAACAGAAAGTGGTCTTTTCTGCTTGA 476

QY 828 ACCCCGAGATCGAGAGATTGGAGCCCGGAAAGAAATGAGAGAGATGGGAGCT 887

DB 477 ACCCCGAGATCGAGAGATTGGAGCCCGGAAAGAAATGAGAGAGATGGGAGCT 536

QY 888 TGACCTGAACGGGACGATTGTTGGTCGCAACCCGGTGAATGCCAAACCGTCCAGCA 947

DB 537 TGAACCTGAACGGGACGATTGTTGGTCGCAACCCGGTGAATGCCAAACCGTCCAGCA 596

QY 948 GGAACCTGAACGGGACGATTGTTGGTGAAGCCGACAAAGTGAAGAGAGCGTGAACATCAT 1007

DB 597 GGAACCTGAACGGGACGATTGTTGGTGAAGCCGACAAAGTGAAGAGAGCGTGAACATCAT 656

QY 1008 CTGCCAGAGACCTGTGTCAGACCTGGCTGAGTTTGGAGCCCTGCATGCACACGTTGGGC 1067

DB 657 CTGCCAGAGACCTGTGTCAGACCTGGCTGAGTTTGGAGCCCTGCATGCACACGTTGGGC 716

QY 1068 GGCCTGTACTCTGGGCTGAATGAAGCGCTGCTGCTGTCTACCTGCGCGTTCGGT 1127

DB 717 GGCCTGTACTCTGGGCTGAATGAAGCGCTGCTGCTGTCTACCTGCGCGTTCGGT 776

QY 1128 GGAGCGATCTGTAAAAACCAATCCGAAACAACCTGGTGAAGAGATCCATCATCAGCA 1187

DB 777 GGAGCGATCTGTAAAAACCAATCCGAAACAACCTGGTGAAGAGATCCATCATCAGCA 836

QY 1188 TCCAGACAAGAGTCGCAATGAAGAAATGTGCAAAAGTATGATGCCAGAAATTAATCAC 1247

DB 837 TCCAGACAAGAGTCGCAATGAAGAAATGTGCAAAAGTATGATGCCAGAAATTAATCAC 896

QY 1248 TCAGACATGCTGTCAGCCCAAGTCAAGCGGCTTTTCTGTATGAAGAAAGGAGTTTCA 1307

DB 897 TCAGACATGCTGTCAGCCCAAGTCAAGCGGCTTTTCTGTATGAAGAAAGGAGTTTCA 956

QY 1308 GGAACCTGTGAAGCTGTGAGACGTTGACAGTGAAGTCCAGACATTAAGCCAGCATAGT 1367

DB 957 GGAACCTGTGAAGCTGTGAGACGTTGACAGTGAAGTCCAGACATTAAGCCAGCATAGT 1016

QY 1368 CGTGTCCGGAAGTCTCTGAGTACAGAAAGGAGCGCGAGCCCTCCCACTGCCAGC 1427

DB 1017 CGTGTCCGGAAGTCTCTGAGTACAGAAAGGAGCGCGAGCCCTCCCACTGCCAGC 1076

QY 1428 ACCCGAGGCGAGCCAGAGAGCCCAAGGAGCCCTGGGGAGTGAACCTTCACGCTGTCAG 1487

DB 1077 ACCCGAGGCGAGCCAGAGAGCCCAAGGAGCCCTGGGGAGTGAACCTTCACGCTGTCAG 1136

QY 1488 CCTGACGACAGAGTCCAGAGATTACGTGTCCTGTGAAGGAAGCAAGCCGCTGTGCAC 1547

DB 1137 CCTGACGACAGAGTCCAGAGATTACGTGTCCTGTGAAGGAAGCAAGCCGCTGTGCAC 1196

QY 1548 CTGCTGCTTCCAGCCCATGCCGACCCGAGAGAGCGAGCGCAGACAGACCCGCTGTGCAC 1607

DB 1197 CTGCTGCTTCCAGCCCATGCCGACCCGAGAGAGCGAGCGCAGACAGACCCGCTGTGCAC 1256

QY 1608 CCTTGAAGAGTGTGGGTCTGCTGTCAGACCTTTTCTGCCACCTGTATCTGGGGCTGCACCCG 1667

DB 1257 CCTTGAAGAGTGTGGGTCTGCTGTCAGACCTTTTCTGCCACCTGTATCTGGGGCTGCACCCG 1316

QY 1668 GACCGGCTGCTACGAGCTGCTGAGCCCGGTTTGTG 1702

DB 1317 GACCGGCTGCTACGAGCTGCTGAGCCCGGTTTGTG 1351

RESULT 12

BD183291

LOCUS BD183291 6235 bp DNA linear PAT 17-JUN-2003

DEFINITION Novel genes and proteins encoded by the genes.

ACCESSION BD183291

VERSION BD183291.1 GI:31875491

KEYWORDS JP 2002345492-A/4.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 6235)

AUTHORS Ohara, O., Nagase, T. and Nakajima, D.

TITLE Novel genes and proteins encoded by the genes

JOURNAL Patent: JP 2002345492-A 4 03-DEC-2002;



COMMENT KAZUSA DNA RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002345492-A/4  
PD 03-DEC-2002  
PF 26-FEB-2002 JP 2002049009  
PI OSMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA  
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,  
A61P25/14,  
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02  
CC Novel genes and proteins encoded by the genes FH Key  
Location/Qualifiers  
FT CDS (50) . . (1414) .  
source 1. . 6235  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 34.7%; Score 930; DB 6; Length 6235;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1160; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1499 CAGTCCAGATTAAGTGTGCTCTGTGCAAGAAAGCCAGCTGTGCACTGTGCTTCC 1558  
DB 4447 CAGTCCAGATTAAGTGTGCTCTGTGCAAGAAAGCCAGCTGTGCACTGTGCTTCC 4506  
QY 1559 AGCCCAATGCCGAGAGCGGAGCGGAGCGAGGAGCCGCGTGTGCGCCCTCAGCAGT 1618  
DB 4507 AGCCCAATGCCGAGAGCGGAGCGGAGCGGAGCGGAGCGGCGTGTGCGCCCTCAGCAGT 4566  
QY 1619 GAGCGGTGCGCGAGCGCTTCTGCGCACTGTATGCGGCGTGCACCCGAGCGGCTGT 1678  
DB 4567 GAGCGGTGCGCGAGCGCTTCTGCGCACTGTATGCGGCGTGCACCCGAGCGGCTGT 4626  
QY 1679 ACCGCTGCTGCGCGCTTGTGTGAGCTCAACCTGTGAGCAAGTGTGTGAGCGGCTGC 1738  
DB 4627 ACCGCTGCTGCGCGCTTGTGTGAGCTCAACCTGTGAGCAAGTGTGTGAGCGGCTGC 4686  
QY 1739 TGAACAACAACAGCTACAGATCAAGATCACTGTGAAGATTAAGTGTGAGCAAGGTTGA 1798  
DB 4687 TGAACAACAACAGCTACAGATCAAGATCACTGTGAAGATTAAGTGTGAGCAAGGTTGA 4746  
QY 1799 CAGTGAAGAAACAGTGTGAGCGGAGCGCTGCGCGAGCGGAGGAGTGTCTGTCT 1858  
DB 4747 CAGTGAAGAAACAGTGTGAGCGGAGCGCTGCGCGAGCGGAGGAGTGTCTGTCTGT 4806  
QY 1859 CTGATTAACAGAGTCAAGGAGAGACCGTGTGTGTACTGTGTGCGCTGCGAGCTTCC 1918  
DB 4807 CTGATTAACAGAGTCAAGGAGAGACCGTGTGTGTACTGTGTGCGCTGCGAGCTTCC 4866  
QY 1919 GTGAGCTGACCTATAGATATGCGAGAAACATTCCTGCTTCCAGATTGCGAGTGCCTGTA 1978  
DB 4867 GTGAGCTGACCTATAGATATGCGAGAAACATTCCTGCTTCCAGATTGCGAGTGCCTGTA 4926  
QY 1979 CATTCCGCTCTGCTGTGTAATGCGGCGCTTAATGCGCGCACTGAGTGAAGCTCACACG 2038  
DB 4927 CATTCCGCTCTGCTGTGTAATGCGGCGCTTAATGCGCGCACTGAGTGAAGCTCACACG 4986  
QY 2039 CCAATGAATTCATCATATCTGTGAACAAGAGTTCANAAACTAAGCATCCAGAGGCC 2098  
DB 4987 CCAATGAATTCATCATATCTGTGAACAAGAGTTCANAAACTAAGCATCCAGAGGCC 5046  
QY 2099 CTGAGAGCTTTTCAAGCACTGAGAGTGAAGAGCGTGTGTTTAAATACAGAGACAAGCA 2158  
DB 5047 CTGAGAGCTTTTCAAGCACTGAGAGTGAAGAGCGTGTGTTTAAATACAGAGACAAGCA 5106  
QY 2159 CGTCAAGGTGTTTCAAGCGCCCTGAGGGAAGGAGCGCGGTCTTCCAGAGGTCTCT 2218  
DB 5107 CGTCAAGGTGTTTCAAGCGCCCTGAGGGAAGGAGCGCGGTCTTCCAGAGGTCTCT 5166  
QY 2219 GGGGTGACTTCTGTGAGAGCTTTTAACTCTGAGTGAAGACCTCCCGAGAGCCCGCGG 2278

DB 5167 GGGGTGACTTCTGTGAGAGCTTTAACCTCTGAGTGAAGACCTCCCGAGAGCCCGCGG 5224  
QY 2279 GGGGTGAGCGCGCGCTCTGTGAGAGCGCTGAGAGGCGCTGTGTGTGCAATCAGACAGCA 2338  
DB 5225 GGGGTGAGCGCGCGCTCTGTGAGAGCGCTGAGAGGCGCTGTGTGTGCAATCAGACAGCA 5284  
QY 2339 GACGAAGCTTTCTGTGAATGAGCGCGCGCTCCCGCAGAGAGGCGCACTTTTGTCTTTGT 2398  
DB 5285 GACGAAGCTTTCTGTGAATGAGCGCGCGCTCCCGCAGAGAGGCGCACTTTTGTCTTTGT 5344  
QY 2399 ACATTTTCCGAAATCAAGTAAAGCAAGTGTGTTTCAAGAAAGTTTCAAGGAGCA 2458  
DB 5345 ACATTTTCCGAAATCAAGTAAAGCAAGTGTGTTTCAAGAAAGTTTCAAGGAGCA 5404  
QY 2459 AGGGCAAGTTTATCAAAAACATTTGTTTCAAGAGAGAGGAGCATTTTACGCTTACG 2518  
DB 5405 AGGGCAAGTTTATCAAAAACATTTGTTTCAAGAGAGAGGAGCATTTTACGCTTACG 5464  
QY 2519 GAGTACACAAATATCTGCTGTGAGAAACACAGCATTTTATTTTATTTTAA 2578  
DB 5465 GAGTACACAAATATCTGCTGTGAGAAACACAGCATTTTATTTTATTTTAA 5524  
QY 2579 TAGGTTGTGCTTATCTTATTAATTAATTAATGTCACAACTGTAGACAAATTAAT 2638  
DB 5525 TAGGTTGTGCTTATCTTATTAATTAATTAATGTCACAACTGTAGACAAATTAAT 5584  
QY 2639 TAATTTATTAATTTACAAATTTGAC 2661  
DB 5585 TAATTTATTAATTTACAAATTTGAC 5607

RESULT 13  
AK090948  
AK090948  
LOCUS  
DEFINITION  
Homo sapiens cDNA FLJ33629 f1s, clone BMY2022454, highly similar  
to Homo sapiens cell cycle checkpoint protein CHFR mRNA.  
AK090948  
VERSION  
AK090948.1 GI:21749208  
KEYWORDS  
oligo capping; f1s (full insert sequence).  
SOURCE  
Homo sapiens  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS  
Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,  
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,  
Sekine, M., Ogiya, M., Nishi, T., Shibahara, T., Tanaka, T.,  
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,  
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Magatsuma, M.,  
Shiratori, A., Sudo, H., Hoshi, T., Kaku, Y., Kodaira, H., Kondo, H.,  
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,  
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,  
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,  
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,  
Hirooka, S., Chiba, Y., Iehida, S., Ono, Y., Takiguchi, S., Watanabe, S.,  
Yoshida, M., Hoshino, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,  
Hara, H., Tanabe, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,  
Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A.,  
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,  
Shiohara, N., Sano, S., Moriya, S., Momiya, H., Satoh, N., Takami, S.,  
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,  
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,  
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,  
Kumagai, A., Iwakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,  
Tashiro, H., Tanigami, A., Fujisawa, T., Ono, T., Yamada, K., Fujii, Y.,  
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,  
Inagaki, H., Ikema, Y., Okamoto, S., Okita, R., Kawakami, T.,  
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,  
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,  
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,  
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,  
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamasita, R.,  
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Ito, T. and Sugano, S.

**TITLE** Complete sequencing and characterization of 21,243 full-length human cDNAs  
**JOURNAL** Nat. Genet. 36 (1), 40-45 (2004)  
**PUBMED** 14702039  
**REFERENCE** 2  
**AUTHORS** Nishimura, K., Matsushima, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Ito, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hiro, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuko, Y., Nagai, K. and Isogai, T.  
**TITLE** UNPUBLISHED  
**JOURNAL** UNPUBLISHED  
**REFERENCE** 3 (bases 1 to 2257)  
**AUTHORS** Isogai, T. and Yamamoto, J.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kametani, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.  
**FEATURES**  
 source  
 1..2257  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="BRAMY2022454"  
 /issue\_type="amygdala"  
 /clone\_id="BRAMY2"  
 /note="Cloning vector: pME18FLJ3"

**ORIGIN**  
 Query Match 33.9%; Score 909; DB 9; Length 2257;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 959; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1702 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 1761  
 669 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 728  
 1762 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 1821  
 729 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 788  
 1822 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 1881  
 789 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 848  
 1882 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 1941  
 849 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 908  
 1942 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 2001  
 909 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 968  
 2002 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 2061  
 969 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 1028  
 2062 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 2121  
 1029 GAGCTCACTGAGTGAAGTGTCTGACGCGGTCTGACACACACAGTACAGTCA 1088

QY 2122 GTGAAGAGAGCGTGTCTTTTAAATACAGAGACAGCAGTCAAGTGTTCACAGCCC 2181  
 DB 1089 GTGAAGAGAGCGTGTCTTTTAAATACAGAGACAGCAGTCAAGTGTTCACAGCCC 1148  
 QY 2182 CTGAGGAGAGAGAGCGAGGCTCTCCGACAGGCTCTGAGGAGTCTTCTGTGAGCTT 2241  
 DB 1149 CTGAGGAGAGAGAGCGAGGCTCTCCGACAGGCTCTGAGGAGTCTTCTGTGAGCTT 1208  
 QY 2242 TTACCTCTGAGTGAAGTGTCTTCCGACAGGCTCTGAGGAGTCTTCTGTGAGCTT 2301  
 DB 1209 TTACCTCTGAGTGAAGTGTCTTCCGACAGGCTCTGAGGAGTCTTCTGTGAGCTT 1268  
 QY 2302 AGCGTGTGAGAGGCTCTGAGTGTGATACAGAGACAGAGACAGAGCTTCTGTGAGCTT 2361  
 DB 1269 AGCGTGTGAGAGGCTCTGAGTGTGATACAGAGACAGAGACAGAGCTTCTGTGAGCTT 1328  
 QY 2362 GCGCGTCTCCGCGAGAGAGGAGGAGTGTCTCTTTTGTACATTTTCCGAACTACAGTTA 2421  
 DB 1329 GCGCGTCTCCGCGAGAGAGGAGGAGTGTCTCTTTTGTACATTTTCCGAACTACAGTTA 1388  
 QY 2422 AGCAGAGTCTGTTTTCAGAGAGAGTTCAGAGAGAGGAGGAGTTCAGAGAGAGTTC 2481  
 DB 1389 AGCAGAGTCTGTTTTCAGAGAGAGTTCAGAGAGAGGAGGAGTTCAGAGAGAGTTC 1448  
 QY 2482 GTTTCAGAGAGAGGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 2541  
 DB 1449 GTTTCAGAGAGAGGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 1508  
 QY 2542 GGGAGAGAGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 2601  
 DB 1509 GGGAGAGAGAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 1568  
 QY 2602 TAAGTTTAAATGTACAACTGTACAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 2661  
 DB 1569 TAAGTTTAAATGTACAACTGTACAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTC 1628

**RESULT 14**  
 AX713476 2186 bp DNA linear PAT 15-APR-2003  
 AX713476  
 LOCUS Sequence 160 from Patent EP1293569.  
 DEFINITION AX713476  
 ACCESSION AX713476  
 VERSION AX713476.1 GI:29888332  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE**  
 1  
**AUTHORS** Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Ito, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuko, Y.  
**TITLE** Full-length cDNAs  
**JOURNAL** Patent: EP 1293569-A 160 19-MAR-2003;  
 Helix Research Institute (HRI); Research Association for Biotechnology (RAB)  
**FEATURES**  
 source  
 1..2186  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

**ORIGIN**  
 Query Match 23.6%; Score 633; DB 6; Length 2186;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 863; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1796 TGACATGAGAGAGAGTGTGACGAGAGCTGTGCTCTTCACGCGGAGTGTTCCTGC 1855  
 DB 695 TGACATGAGAGAGAGTGTGACGAGAGCTGTGCTCTTCACGCGGAGTGTTCCTGC 754  
 QY 1856 TGCTGATTTACAGAGTACGCGAGAGACACCGTCTGTGTCTGTGTGCGCTGCGAGCT 1915



QY 1856 TGTCTGATTTACAGAGTCAGGAGACCCGTTCTGTGTTACTGCTGTGGCCCTGGAGCT 1915  
 DB 755 TGTCTATTTACAGAGTCAGGAGACCCGTTCTGTGTTACTGCTGTGGCCCTGGAGCT 814  
 QY 1916 TCCGTGAGCTGACCTTACAGTATCGGCAGAACATTCCTGCTTCCAGTTGCCAGTGGCCG 1975  
 DB 815 TCCGTGAGCTGACCTTACAGTATCGGCAGAACATTCCTGCTTCCAGTTGCCAGTGGCCG 874  
 QY 1976 TAAATCCCGTCTGACTGCTGACTGCGGCGCTGAACTGCGCACTCAGGTGAAGTCACT 2035  
 DB 875 TAAATCCCGTCTGACTGCTGACTGCGGCGCTGAACTGCGCACTCAGGTGAAGTCACT 934  
 QY 2036 AGCCCATGAATTCATCATATCTGTGAACAGCAAGGTTCAAAACTAAGACATCCAGAG 2095  
 DB 935 AGCCCATGAATTCATCATATCTGTGAACAGCAAGGTTCAAAACTAAGACATCCAGAG 994  
 QY 2096 GGCCTGAGAGCTTTACAGACTGAGAGTGAAGAGAGCTGTTTAAATACAGAGACAA 2155  
 DB 995 GGCCTGAGAGCTTTACAGACTGAGAGTGAAGAGAGCTGTTTAAATACAGAGACAA 1054  
 QY 2156 GCACTGAAAGTGTTCACAGCCCTGAGAGGAGGACGAGGCTCCGACAGTGC 2215  
 DB 1055 ACACTGAAAGTGTTCACAGCCCTGAGAGGAGGACGAGGCTCCGACAGTGC 1114  
 QY 2216 TCTGGGGTGAATCTCTGTGAGAGCTTTTAACTGAGTGAAGCCCTCCGAGAGCC 2275  
 DB 1115 TCTGGGGTGAATCTCTGTGAGAGCTTTTAACTGAGTGAAGCCCTCCGAGAGCC 1172  
 QY 2276 GGGGGCCGAGCCGCTCTGTGAGAGCTGAGAGGCTGCTGTGAGTGAAGCTC 2335  
 DB 1173 GGGGGCCGAGCCGCTCTGTGAGAGCTGAGAGGCTGCTGTGAGTGAAGCTC 1232  
 QY 2336 AAGAGAGAGCTTTCTGTGAACATGCGGCGCTCCGACAGAGGAGCTTTGCTCTT 2395  
 DB 1233 AAGAGAGAGCTTTCTGTGAACATGCGGCGCTCCGACAGAGGAGCTTTGCTCTT 1292  
 QY 2396 TGAACATTTCCGAAACCTACAGTTAAGAGAGAGTGTGTTTCAAGAAAGTTTCAAGG 2455  
 DB 1293 TGAACATTTCCGAAACCTACAGTTAAGAGAGAGTGTGTTTCAAGAAAGTTTCAAGG 1352  
 QY 2456 AAGAGAGAGCTTTTCAAAAATTTGTTTCAAGAGAGAGAGTAACTTACAGCTTA 2515  
 DB 1353 AAGAGAGAGCTTTTCAAAAATTTGTTTCAAGAGAGAGAGTAACTTACAGCTTA 1412  
 QY 2516 CAGGAGCTACACATATCTGCTGCTGAGAAACCAAGCATTTTATCTATTTTATTT 2575  
 DB 1413 CAGGAGCTACACATATCTGCTGCTGAGAAACCAAGCATTTTATCTATTTTATTT 1472  
 QY 2576 TAAATGCTTGTGCTTATCTTCTAATAGATTAAATGTCACAACTGACACAAATA 2635  
 DB 1473 TAAATGCTTGTGCTTATCTTCTAATAGATTAAATGTCACAACTGACACAAATA 1532  
 QY 2636 AATAATTTATATTTTACAAATTTGAC 2661  
 DB 1533 AATAATTTATATTTTACAAATTTGAC 1558

RESULT 16  
 LOCUS CQ719982 1085 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 5916 from Patent WO02068579.  
 ACCESSION CQ719982  
 VERSION CQ719982.1 GI:42280839  
 KEYWORDS  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 TITLE Kite, such as nucleic acid arrays, comprising a majority of  
 humenexons or transcripts, for detecting expression and other uses  
 thereof

JOURNAL Patent: WO 02068579-A 5916 06-SEP-2002;  
 FEATURES  
 source  
 1..1085  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 18.0%; Score 483; DB 6; Length 1085;  
 Best Local Similarity 100.0%; Pred. No. 4,1e-268;  
 Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 710 GGTCTGGGGGTGTGTGGCATCTCCCTTAAAGAGAGTGTCTCTGTGGCAAGTGAAG 769  
 DB 1 GGTCTGGGGGTGTGTGGCATCTCCCTTAAAGAGAGTGTCTCTGTGGCAAGTGAAG 60  
 QY 770 TCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGTGCCTTTGCTGTGGAAC 829  
 DB 61 TCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGTGTCTCTGTGTGGAAC 120  
 QY 830 CCCAGATCAGAGAGATTTGAGCCCTGGAAGAAATGAGAGAGATGGGACCTTG 889  
 DB 121 CCCAGATCAGAGAGATTTGAGCCCTGGAAGAAATGAGAGAGATGGGACCTTG 180  
 QY 890 ACTTGAACGGGAGATTTGTTGTTGCTGCAACACCGCTGAAATGCCAACCTGCCAGAG 949  
 DB 181 ACTTGAACGGGAGATTTGTTGTTGCTGCAACACCGCTGAAATGCCAACCTGCCAGAG 240  
 QY 950 AGCTGAGAGAGCGGCTGGGAGAGCAGACAAAGATGAGAGAGCGCTGACATCATCT 1009  
 DB 241 AGCTGAGAGAGCGGCTGGGAGAGCAGACAAAGATGAGAGAGCGCTGACATCATCT 300  
 QY 1010 GCCAGACCTGCTGACAGCATCTGCTGAGTTTGCAGCCCTGATGACAGAGTTCTGCGCG 1069  
 DB 301 GCCAGACCTGCTGACAGCATCTGCTGAGTTTGCAGCCCTGATGACAGAGTTCTGCGCG 360  
 QY 1070 CTTCCTACTCCGGCTGATGAGAGCGCTGCTGCTGTCTTACCTGCGCGCTGCTGCG 1129  
 DB 361 CTTCCTACTCCGGCTGATGAGAGCGCTGCTGCTGTCTTACCTGCGCGCTGCTGCG 420  
 QY 1130 AGCGATCTGTAAAAACCACTCTCAACAACTGCTGGAAGCATACCTCATCCAGATC 1189  
 DB 421 AGCGATCTGTAAAAACCACTCTCAACAACTGCTGGAAGCATACCTCATCCAGATC 480  
 QY 1190 CAG 1192  
 DB 481 CAG 483

RESULT 17  
 LOCUS AX868758 816 bp DNA linear PAT 17-DEC-2003  
 DEFINITION Sequence 3663 from Patent EP1074617.  
 ACCESSION AX868758  
 VERSION AX868758.1 GI:40023621  
 KEYWORDS  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Ota, T., Iwagai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
 TITLE Primers for synthesizing full-length cDNA and their use  
 Patent: EP 1074617-A 3663 07-FEB-2001;  
 JOURNAL Research Association for Biotechnology (Jp)  
 FEATURES  
 source  
 1..816  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN



OY	87	CCCGATGGAACGACCCCGAAGAAAGGCAAGAGATGCGCGCGCGACGCTTGCGGAGACGGCT	146
Db	61	CCCGATGGAACGACCCCGAAGAAAGGCAAGAGATGCGCGCGCGACGCTTGCGGAGACGGCT	120
OY	147	CCTGCGCTCTGGGCGCGGAGAGAGGCGAGCGACGTCCTCTTGAGGAAACCGGAGATGAGAC	206
Db	121	CCTGCGCTCTGGGCGCGGAGAGAGGCGAGCGACGTCCTCTTGAGGAAACCGGAGATGAGAC	180
OY	207	CATCGGGCGGAGACGAGGCTTGGCACTTTCTTCCCGACCAATAACTGCTCTTGAGAA	266
Db	181	CATCGGGCGGAGACGAGGCTTGGCACTTTCTTCCCGACCAATAACTGCTCTTGAGAA	240
OY	267	TCACGTGTGAATTTGTAGTGAATGAAAAATCAGGTGAGTGAACATCGGAAGATTACCGACAC	326
Db	241	TCACGTGTGAATTTGTAGTGAATGAAAAATCAGGTGAGTGAACATCGGAAGATTACCGACAC	300
OY	327	CAGTGAACAGTGATTAAACAGCTGAAGGTGTTAAGAAAGCAGACATGCTTTACAGAC	386
Db	301	CAGTGAACAGTGATTAAACAGCTGAAGGTGTTAAGAAAGCAGACATGCTTTACAGAC	360
OY	387	TGGGGATGTCACTACTTGTGTACAGGAAAGATGAACCGAACAACACGTGGCATACT	446
Db	361	TGGGGATGTCACTACTTGTGTACAGGAAAGATGAACCGAACAACACGTGGCATACT	420
OY	447	CTATGAATCTTTAAGTGAAGAAAGCAAGGCGATGACACAAGAAATCCTTTGA	494
Db	421	CTATGAATCTTTAAGTGAAGAAAGCAAGGCGATGACACAAGAAATCCTTTGA	468

RESULT	20
LOCUS	BD148838
DEFINITION	BD148838 824 bp DNA linear PAT 17-JAN-2002
ACCESSION	BD148838
VERSION	BD148838.1 GI:27854596
KEYWORDS	JP 2002191363-A/3681.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (base) 1 to 824
REFERENCE	Ota,T., Iwoga,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakematsu,A., Nagai,K. and Otsuki,T. Primer for synthesizing full-length cDNA and use thereof Patent: JP 2002191363-A 3681 09-JUL-2002;
JOURNAL	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/3681
COMMENT	PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, PC C12P21/02,C12O1/68//C12P21/08,G06F17/30 C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT source 1..824 Location/Qualifiers 1..824 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
FEATURES	source
ORIGIN	Query Match 17.5%; Score 468; DB 6; Length 824; Best Local Similarity 100.0%; Pred. NO. 2.le-259; Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 21	AX873789/c	518 bp	DNA	linear	PAT 17-DEC-2003
LOCUS	AX873789	518 bp	DNA	linear	PAT 17-DEC-2003
DEFINITION	Sequence 8694 from Patent EP1074617.				
ACCESSION	AX873789				
VERSION	AX873789.1	GI:40028580			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE					
1					
AUTHORS	Ora T., Isogai T., Nishikawa T., Hayashi K., Saito K., Yamamoto J.,				
	Ishii S., Sugiyama T., Wakamatsu A., Nagai K. and Otsuki T.				
TITLE	Primers for synthesizing full-length cDNA and their use				
JOURNAL	Patent: EP 1074617-A 8694 07-FEB-2001;				
	Research Association for Biotechnology (JP)				
FEATURES					
source	location/Qualifiers				
	1..518				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	15.9%; Score 426; DB 6; Length 518;				
Best Local Similarity	99.8%; Pred. No. 5.2e-25;				
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
OY	2185 AGGAGAGGACGACGAGGCTCCGACAGGTGCTCGGGGTGACTCTTCTGTGAGAGCTTTT	2244			
DB	503 AGGAGAGGACGACGAGGCTCCGACAGGTGCTCGGGGTGACTCTTCTGTGAGAGCTTTT	444			
OY	2245 ACCCTCTGAGTGAGACCTCCGCCAGAGCCCGGGGGCCGACGCCCTCTCGGTGAGC	2304			
DB	443 ACCCTCTGAGTGAGACCTCCGCCAGAGCCCGGGGGCCGACGCCCTCTCGGTGAGC	384			
OY	2305 GCTGGGACGAGGCTGTGTGTGGCATCAGACGACGACGAGAGCTTTCTGTAAATGCGGC	2366			
DB	383 GCTGGGACGAGGCTGTGTGTGGCATCAGACGACGACGAGAGAGCTTTCTGTAAATGCGGC	324			



QY	2365	CGTCCCGCCGAAAGGGGCGAGTTTGCTCTTTGTGCATTTTCCGAACTACAGTTAAAGC	2422
DB	323	CGTCCCGCCGAAAGGGGCGAGTTTGCTCTTTGTGCATTTTCCGAACTACAGTTAAAGC	264
QY	2425	AGAACTCTGTTTTTCAGAAAAGATTTCMAAGGGAAGGGCGAACTTTATCAAAAACATTGTT	2484
DB	263	AAAATCTGTTTTTCAGAAAAGATTTCMAAGGGAAGGGCGAACTTTATCAAAAACATTGTT	204
QY	2485	TCAGAGAAAGGAGACATAAGTTTACAGCTTACAGACGTACACATATCTCTGCTGTGGG	2544
DB	203	TCAGAGAAAGGAGACATAAGTTTACAGCTTACAGACGTACACATATCTCTGCTGTGGG	144
QY	2545	AAAACCAAGCAGCTTTTATCTATTTTTTTTAAATAGTTGGTGGTCTATCTTCAATTA	2604
DB	143	AAAACCAAGCAGCTTTTATCTATTTTTTTTAAATAGTTGGTGGTCTATCTTCAATTA	84
QY	2605	GATTTAAATGTCACAAACTGTAGCAACAATAATTAATTTAATTTACAAATTGAC	2661
DB	83	GATTTAAATGTCACAAACTGTAGCAACAATAATTAATTTAATTTAATTTCAATTGAC	27
RESULT 22			
LOCUS	BD153851	518 bp	DNA
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		linear PAT 17-JAN-2003
ACCESSION	BD153851		
VERSION	BD153851.1	GI:27859609	
KEYWORDS	JP 2002191363-N/8694.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 518)		
TITLE	Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., and Otsuki,T.		
JOURNAL	Ishii,S., Sugiyama,T., Wakematsu,A., Nagai,K. and Otsuki,T.		
COMMENT	Primer for synthesizing full-length cDNA and use thereof		
	Patent: JP 2002191363-A 8694 09-JUL-2002;		
	HELIX RESEARCH INSTITUTE		
	OS Homo sapiens (human)		
	PN JP 2002191363-A/8694		
	PD 09-JUL-2002		
	PF 28-JUL-2000 JP 2000280990		
	PI TOSHIO OYA,TAKAO ISEGAI,IETSUO NISHIKAWA,KOJI HAYASHI,KAORU		
	PI SAITO,		
	PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,		
	PI KEIICHI NAGAI,IETSUJI OTSUKI		
	PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC		
	10,		
	PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC		
	Primer for synthesizing full-length cDNA and use thereof FH Key		
FEATURES	Location/Qualifiers		
source	FT source	1..518	
	Location/Qualifiers		
	1..518		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	15.9%	Score 426;	DB 6;
Best Local Similarity	99.8%	Pred. No. 5.2e-235;	
Matches	476;	Conservative 0;	Mismatches 1;
		Indels 0;	Gaps 0;
QY	2185	AGGGAAGGAGCGCAGAGGTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGCTTTT	2244
DB	503	AGGGAAGGAGCGCAGAGGTCTCCGACAGGTGCTCTGGGGTGACTCTTCTGTGAGCTTTT	444
QY	2245	ACCTCTGTAGTGAACCTCTCCCAAGCCCGGGGGCGCGACCGCCCTCTGGTGAAC	2304
DB	443	ACCTCTGTAGTGAACCTCTCCCAAGCCCGGGGGCGCGACCGCCCTCTGGTGAAC	384

QY	2305	GTGGGGCAGGCGCTCGTGTGGCATTCAGAGCAGAGACCAACGCTTTCTGTATCAATGCGGC	2305
Db	383	GTGGGGCAGGCGCTCGTGTGGCATTCAGAGCAGAGACCAACGCTTTCTGTATCAATGCGGC	324
QY	2365	CGTCCCGCCAGAGAGGCGAGGTTTGTCTCTTTGTGTACATTTTCGAAACTACAGTTAAAGC	2424
Db	323	CGTCCCGCCAGAGAGGCGAGGTTTGTCTCTTTGTGTACATTTTCGAAACTACAGTTAAAGC	264
QY	2425	AGAACTCTGTTTTCAAGAAAAGTTTTCAGAGGAGAAAGGCAAGTTTATCAAAAACATGTT	2484
Db	263	AAAGTCTGTTTTCAAGAAAAGTTTTCAGAGGAGAAAGGCAAGTTTATCAAAAACATGTT	204
QY	2485	TCAGAGGAAGGAGCATATGTTTACAGCGCTACAGAGCTACACATATCCTGCTGTGGG	2544
Db	203	TCAGAGGAAGGAGCATATGTTTACAGCGCTACAGAGCTACACATATCCTGCTGTGGG	144
QY	2545	AAACACAGACATTTTATCTATTTTATTTTAAATAGGTTGGTCTTATCTTAAATTA	2604
Db	143	AAACACAGACATTTTATCTATTTTATTTTAAATAGGTTGGTCTTATCTTAAATTA	84
QY	2605	GATTTAATGTCACAAACTGTACGACCAATATATTTATTTATTTTACAAATGAC	2661
Db	83	GATTTAATGTCACAAACTGTACGACCAATATATTTATTTATTTTACAAATGAC	27
RESULT 23			
AC127070			
LOCUS	AC127070	181012 bp	DNA linear
DEFINITION	Human sapiens 12 BAC Rpl1-68H1 (Roswell Park Cancer Institute Human BAC Library) complete sequence.		
AC127070			
ACCESSION	AC127070.10	GI:23343648	
KEYWORDS	HTC		
SOURCE	Human sapiens (human)		
ORGANISM	Human sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 181012)		
	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alldrooks, S.L., Amaratunga, H.C., Are, J.R., Asyle, M., Bakes, T., Barbarta, J., Benton, J., Bimaga, K., Blanchburg, K., Bonnah, D., Bowls, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Bulay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.P., Carter, M., Cavazos S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, Cleveland, C.D., Cox, C., Coyle, M.D., Dacton, S.R., David, R., Delaney, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H., Douthett, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Eamthair, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlik, P., Haws, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshihes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, W., Loubege, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Me, J., Maheshwari, M., Mapua, P., Marcondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Maunley, E., McLeod, M.P., Meador, M., Mei, G., Merche, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Monabhat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Orzuyen, N., Nickerson, E., Nwokweto, S., Ognih, M., Okumou, G., Orzuyen, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Qulles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saver, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoostari, N., Sisson, I.,		





misc\_feature clone\_end:SP6"  
103635..181438  
/note="assembly\_name:Contig21  
clone\_end:T7"

## ORIGIN

Query Match 14.4%; Score 386; DB 2; Length 181438;  
Best Local Similarity 99.5%; Pred. No. 2e-211;  
Matches 616; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

2043 GAAATTCATCATATCTGTGAACAGACAAAGTTCAAAAATCAAGATCCAGAGGCCCTTA 2102  
108641 GAAATTCATCATATCTGTGAACAGACAAAGTTCAAAAATCAAGATCCAGAGGCCCTTA 108700  
2103 GCAGCTTTCAGACCTGAGAGTGAAGAGCGTGTATTTAAATACAGACAGACAGCCTC 2162  
108701 GCAGCTTTCAGACCTGAGAGTGAAGAGCGTGTATTTAAATACAGACAGACAGCCTC 108760  
2163 AAGGTCTTTCACAGCCCTTGAAGAGAGAGCGGTCTCCACAGAGTGTCTGAGG 2222  
108761 AAGGTCTTTCACAGCCCTTGAAGAGAGAGCGGTCTCTCCACAGAGTGTCTGAGG 108820  
2223 TGACCTTCTGTGAGAGCTTTTACCTCTGAGTGAAGACCTGCCAGAGCCCGGAGGAGC 2282  
108821 TGACCTTCTGTGAGAGCTTTTACCTCTGAGTGAAGACCTGCCAGAGCCCGGAGGAGC 108878  
2283 GCAGCCCGGAGGAGTGTGAGAGCGTGTGAGAGCGGTCTGAGTGAAGAGCAGAGAGC 2342  
108879 GCAGCCCGGAGGAGTGTGAGAGCGGTGTGAGAGCGGTCTGAGTGAAGAGCAGAGAGC 108938  
2343 AAGCCCTTCTGTGAACATGAGGAGCGTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGC 2402  
108939 AAGCCCTTCTGTGAACATGAGGAGCGTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGC 108998  
2403 TTTCGCAAACTACAGTTAAAGAGAGAGTCTGTTTTCAGAGAAAAGTTTCAAGGAGAGAGG 2462  
108999 TTTCGCAAACTACAGTTAAAGAGAGAGTCTGTTTTCAGAGAAAAGTTTCAAGGAGAGAGG 109058  
2463 CAAGTTTATCAAAAACATTTGTTTCAGAGAGAGGAGAGCAGTTTACAGCTTACAGAGAGC 2522  
109059 CAAGTTTATCAAAAACATTTGTTTCAGAGAGAGGAGAGCAGTTTACAGCTTACAGAGAGC 109118  
2523 TACACAAATATCTGCTGCTGAGGAGAAACACAGAGATTTTATCTATTTTATTTATATAG 2582  
109119 TACACAAATATCTGCTGCTGAGGAGAAACACAGAGATTTTATCTATTTTATTTATATAG 109178  
2583 TTTCGCTTATCTTCTATATTAAGATTTAATGTCACAACTGTAGACAAATATATATAT 2642  
109179 TTTCGCTTATCTTCTATATTAAGATTTAATGTCACAACTGTAGACAAATATATATAT 109238  
2643 TTTATATTTTACAAATTTGAC 2661  
109239 TTTATATTTTACAAATTTGAC 109257

RESULT 25  
AC023047 181438 bp DNA linear HTG 05-MAR-2001  
LOCUS Homo sapiens chromosome 12 clone RP11-46H11, WORKING DRAFT  
DEFINITION AC023047  
SEQUENCE 6 unordered pieces.  
AC023047.19 GI:9857509  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1. (bases 1 to 181438)  
Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,  
Federpspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,  
Mao,J., Lam,B., Marathe,R., Miranda,M., Morehouse,A.J., Nguyen,M.,  
Oefner,P., Palm,C.J., Ramirez,D., Southwick,A.M., Wilhelmy,J.,  
Yu,S. and Davis,R.W.

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 181438)  
Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D., Federpspiel,N.,  
Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,  
Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,  
Yu,S. and Davis,R.W.

TITLE  
JOURNAL

Direct Submission  
Submitted (07-FEB-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA

## COMMENT

On Aug 19, 2000 this sequence version replaced gi:9665109.

## Genome Center

Center: Stanford DNA Sequencing and Technology Development

## Center code: SDSTDC

Web site: <http://sequence-www.stanford.edu/group/human/>

Contact: hum-info@sequence.stanford.edu

## Project Information

Center project name: 840

Center clone name: RP11-46H11

Sequencing Vector: M13mp18; X02513

Chemistry: Dye-primers; 1% of reads

Chemistry: Dye-terminator Big Dye; 99% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 178770 bases at least Q40

Consensus quality: 179976 bases at least Q30

Consensus quality: 180370 bases at least Q20

Insert size: 16471; agarose-gel

Insert size: 180938; sum-of-coverage

Quality coverage: 10.7x in Q20 bases; agarose-gel

Quality coverage: 9.7x in Q20 bases; sum-of-coverage.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 6 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 7009: contig of 7009 bp in length

\* 7010 7109: gap of unknown length

\* 7110 23344: contig of 16235 bp in length

\* 23345 23445: gap of unknown length

\* 23445 38890: contig of 15446 bp in length

\* 38891 38990: gap of unknown length

\* 38991 55546: contig of 16556 bp in length

\* 55547 103534: gap of unknown length

\* 103535 103634: contig of 47888 bp in length

\* 103635 181438: gap of unknown length

\* 181438: contig of 77804 bp in length.

\* Location/Qualifiers

1. 181438

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="12"

/clone="RP11-46H11"

/clone\_lib="RPCI human BAC library 11"

1. 7009

/note="assembly\_name:Contig16"

7110. 23344

/note="assembly\_name:Contig17"

23445. 38890

/note="assembly\_name:Contig18"

38991. 55546

/note="assembly\_name:Contig19"

55647. 103534

/note="assembly\_name:Contig20

103635. 181438

/note="assembly\_name:Contig21

clone\_end:T7"

## ORIGIN

Query Match 8.1%; Score 218; DB 2; Length 181438;  
Best Local Similarity 100.0%; Pred. No. 1e-113;  
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 493 GAAGCTAACAGAGAAATGTTTCATGGAGACCAAGATACCTCAGGTGACAGGTGACAGG 552  
|||||  
DB 13993 GAAGCTAACAGAGAAATGTTTCATGGAGACCAAGATACCTCAGGTGACAGGTGACAGG 13934  
|||||  
QY 553 CGAGGGGCGGATCCCGGGTCCCTCCGTGTCGCGCCCACTCAGAGTGTGCTTTGAGAA 612  
13933 CGAGGGGCGGATCCCGGGTCCCTCCGTGTCGCGCCCACTCAGAGTGTGCTTTGAGAA 13874  
|||||  
DB 613 CCACAGCATCAATGATGAGGTGACAGACCTTTCCTCCACAGCTTGGCTTTGACGAG 672  
13873 CCACAGCATCAATGATGAGGTGACAGACCTTTCCTCCACAGCTTGGCTTTGACGAG 13814  
|||||  
QY 673 CCTTCTCTGACGAGGCGAGAGCGTTCTCCAGTTGTGG 710  
13813 CCTTCTCTGACGAGGCGAGAGCGTTCTCCAGTTGTGG 13776  
|||||

RESULT 26  
BV180578/c 756 bp DNA linear STS 10-JUN-2004  
LOCUS BV180578 Human DNA (Sequencem) Homo sapiens STS genomic, sequence  
DEFINITION BV180578  
tagged site.  
ACCESSION BV180578  
VERSION BV180578.1 GI:48017719  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Nelson,R.M., Marnellios,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
1 (bases 1 to 756)  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions

TITLE  
JOURNAL  
COMMENT

Genome Res. (2004) In press

Contact: Andreas Braun  
Pharmaceuticals division  
Sequencem, Inc.

3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequencem.com

Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 756.

FEATURES  
source location/Qualifiers

1..756  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequencem)"  
<1..>756

## ORIGIN

Query Match 3.7%; Score 98; DB 11; Length 756;  
Best Local Similarity 100.0%; Pred. No. 3e-44;  
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2564 TATTTTATTTTAAATGTTGGTCTTATCTTATATAGATTAAATGTCACAAACT 2623  
|||||  
DB 726 TATTTTATTTTAAATGTTGGTCTTATCTTATATAGATTAAATGTCACAAACT 667  
|||||  
QY 2624 GTAGACAAAT 2661  
|||||  
DB 666 GTAGACAAAT 629  
|||||

## RESULT 27

BV168191/c 601 bp DNA linear STS 10-JUN-2004  
LOCUS BV168191 Human DNA (Sequencem) Homo sapiens STS genomic, sequence  
DEFINITION BV168191  
tagged site.

ACCESSION BV168191 GI:48001924  
VERSION BV168191.1  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Nelson,R.M., Marnellios,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
1 (bases 1 to 601)  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions

TITLE  
JOURNAL  
COMMENT

Genome Res. (2004) In press

Contact: Andreas Braun  
Pharmaceuticals division  
Sequencem, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequencem.com

Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 601.

FEATURES  
source location/Qualifiers

1..601  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequencem)"  
<1..>601

## ORIGIN

Query Match 3.6%; Score 97; DB 11; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.1e-43;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2565 ATTTTATTTTAAATGTTGGTCTTATCTTATATAGATTAAATGTCACAACTG 2624  
|||||  
DB 601 ATTTTATTTTAAATGTTGGTCTTATCTTATATAGATTAAATGTCACAACTG 542  
|||||  
QY 2625 TAGCACAAAT 2661  
|||||  
DB 541 TAGCACAAAT 505  
|||||

## RESULT 28

AX385952/c 617 bp DNA linear PAT 19-MAR-2002  
LOCUS AX385952 Sequence 880 from Patent WO0214500.  
DEFINITION AX385952  
ACCESSION AX385952  
VERSION AX385952.1 GI:19579082  
KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,  
1 Randazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and  
Labat,I.  
Human genes and gene expression products  
Patent: WO 0214500-A 880 21-FEB-2002;  
CHIRON CORPORATION (US); Hysq Inc. (US)

Location/Qualifiers  
1..617  
/organism="Homo sapiens"

## FEATURES

source

```

/mot_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      2.9%; Score 78; DB 6; Length 617;
Best Local Similarity 100.0%; Pred. No. 1.3e-32;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1966 CCAAGTGGCCGTAACATCCGCTGCTGCTAGTGGGCGGTAACCTGCGCACTCAGGATG 2025
DB 385 CCAAGTGGCCGTAACATCCGCTGCTGCTAGTGGGCGGTAACCTGCGCACTCAGGATG 326
QY 2026 AAGCTCACACGCGCATG 2043
DB 325 AAGCTCACACGCGCATG 308

RESULT 29
LOCUS CQ536585 60 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 6220 from Patent WO0210449.
ACCESSION CQ536585
VERSION CQ536585.1 GI:41502849
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Shoshan,A., Maeseman,A., Mintz,E., Mintz,L. and Faigler,S.
AUTHORS Shoshan,A., Maeseman,A., Mintz,E., Mintz,L. and Faigler,S.
TITLE Oligonucleotide library for detecting rna transcripts and splice
JOURNAL variants that populate a transcriptome
FEATURES Patent: WO 0210449-A 6220 07-FEB-2002;
source CompuGen Inc. (US)
LOCATION/Qualifiers
source 1. 60
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      2.2%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-22;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1843 GGAAGTCTTCTGCTGCTGATTAACAGATCAGGAGACACCGTCTGTTACTGCTGT 1902
DB 1 GGAAGTCTTCTGCTGCTGATTAACAGATCAGGAGACACCGTCTGTTACTGCTGT 60

RESULT 30
LOCUS AC118260 158897 bp DNA linear HTG 28-FEB-2003
DEFINITION Mus musculus clone RP24-374B14, WORKING DRAFT SEQUENCE.
ACCESSION AC118260
VERSION AC118260.3 GI:28604085
KEYWORDS HTG: HTGS PHASE2; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Birren,B., Nuebaum,C. and Lander,E.
JOURNAL Mus musculus, clone RP24-374B14
AUTHORS Unpublished
2 (bases 1 to 158897)
Birren,B., Linton,L., Nuebaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Charato,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

```

```

Hagos,B., Horton,L., Hulme,W., Iley,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Lander,E., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,T., Margis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Melnick,J., Menes,I.,
Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riback,M., Riley,R., Riese,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-Apr-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 158897)
Birren,B., Nuebaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Bouhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collimore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Haggopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iley,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,T., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Menes,I., Mihova,T., Mieng,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,V., Raymond,C., Retta,R., Riese,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 28, 2003 this sequence version replaced gi:28394919.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 374.B.14
Center clone name: L24782
Summary Statistics
Sequencing vector: plasmid, n/a, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.960731
Consensus quality: 158848 bases at least Q40
Consensus quality: 158895 bases at least Q30
Consensus quality: 158897 bases at least Q20
Insert size: 158000; agarose-fp
Insert size: 158897; sum-of-contigs
Quality coverage: 12.6 in Q20 bases; agarose-fp
Quality coverage: 12.6 in Q20 bases; sum-of-contigs
NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced

```



\* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 1 158897: contig of 158897 bp in length.

FEATURES  
 source  
 Location/Qualifiers  
 1. 158897  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP24-374B14"  
 /clone\_lib="RPC1-24 Male Mouse BAC"

## ORIGIN

Query Match 1.4%; Score 38; DB 2; Length 158897;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-09;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1975 GTAACTCCCTCTGCTGCTACTGAGCGCGTAACTG 2012

Db 40304 GTAACTCCCTCTGCTGCTACTGAGCGCGTAACTG 40341

RESULT 31  
 BC049792 3194 bp mRNA linear ROD 30-JUN-2004  
 LOCUS BC049792  
 DEFINITION Mus musculus checkpoint with forkhead and ring finger domains, mRNA  
 (CDNA clone MGC:59419 IMAGE:6510480), complete cds.  
 ACCESSION BC049792.1 GI:29437110  
 VERSION BC049792.1  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3194)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schlier, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.T., Wang, J., Heien, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Scapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Ueding, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mulich, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, X., Gibbs, R.A.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Gitlwood, J., Schmutz, J., Myers, R.M.,  
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
 PUBLISHED  
 12477932  
 2 (bases 1 to 3194)  
 Strausberg, R.  
 Direct Submission  
 Submitted (31-MAR-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)  
 Tissue Procurement: Dr. David Rowe  
 cDNA Library Preparation: Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth  
 Featherstone, Malachi Griffith, Oki Griffith, Ran Guin, Nancy Liao,  
 Kim Macdonald, Amara Mason, Mike R. Mayo, Josh Moran, Ryan Morin,  
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lila Prabhu,  
 Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska,  
 Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacqie  
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clome distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
 Series: IRK Plate: 109 Row: 1 Column: 17  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 27370049.

FEATURES  
 source  
 Location/Qualifiers  
 1. 3194  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="MGC:59419 IMAGE:6510480"  
 /issue\_type="limb, mouse, day 10.5 and 11.5, pool of  
 undifferentiated limb containing undifferentiated limb  
 mesenchyme and early condensing mesenchyme."  
 /clone\_lib="NIH MGC\_134"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6.1"

gene  
 /gene="Cnfr"  
 /note="synonyms: RNF116, C230082M18"  
 /db\_xref="LOCUSID:231600"  
 /db\_xref="MGI:2444898"

## CDS

71..2065  
 /gene="Cnfr"  
 /codon\_start=1  
 /product="checkpoint with forkhead and ring finger  
 domains"  
 /protein\_id="AAH49792.1"  
 /db\_xref="GI:29437111"  
 /db\_xref="LOCUSID:231600"  
 /db\_xref="MGI:2444898"  
 /translation="MEHGEQPPPEPPEWGLRLVGADEDEPQILLMREWTGRR  
 GIDSPSKVLVSGHCKLTVDEISGEVTEPTSTVINKLVKQVPLQSGDI  
 IIVLVKNEPENNVAIVLSLGGKSLTQDSLEANKENNEHYTKDSCGQGDPCVP  
 ILSVPAQTLERQPSRTSDILPASTSTPELTSAQKSSSGSPGNTISPKGR  
 SIVAVGSLSSVYFQDKASFSLSKQHELEPAKRMKGDGLDNLQVLSQ  
 RNAQGVSSSDVDASVDPKMEETLTCILQDLNDVSLQPCMTFCACYSQNER  
 SSICPTCRCPVPRICKNHILNIVKAYLQHPDKSSEEDVSMARNTITDMLQPK  
 VRRSFDEGSEEDLLESDVSESDISQPIYVCGQCEYRQAVQSAPCVSESEL  
 GTALVIGERAPSTASLPTAAPVWCPIQSHAICTCCPQPPDRAAREQDPRVAPQ  
 QCAVCIQPCCHTYNGCTRTGCGCLAPFELMDKCLDGVANNYBEDIKNYIAT  
 RGLTWKSVITSELALQKGVFMISDRITGNTVLCYCGALRSFRELTYQRYQVTPASE  
 LPVTYSRPDCYWGRCRQYVAHAKMFNHI CEQTRFKN"

## ORIGIN

Query Match 1.4%; Score 37; DB 10; Length 3194;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 TTCCGAGCAATTAAGTGTCTCTGAGATCACTGTA 274  
 Db 218 TTCCGAGCAATTAAGTGTCTCTGAGATCACTGTA 254

## RESULT 32

AC087142 321708 bp DNA linear HTG 09-DEC-2000.  
 LOCUS AC087142  
 DEFINITION Mus musculus clone RP23-211L19, WORKING DRAFT SEQUENCE, 65  
 unordered pieces.

ACCESSION AC087142.1 GI:11610868  
 VERSION AC087142.1  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE DOE Joint Genome Institute.  
JOURNAL Sequencing of Mouse  
AUTHORS Unpublished  
REFERENCE 2 (bases 1 to 321708)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (09-DEC-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 1818179  
Center clone name: RPCI-23\_211119  
-----  
Summary Statistics  
Consensus quality: 271436 bases at least Q40  
Consensus quality: 288232 bases at least Q30  
Consensus quality: 297606 bases at least Q20  
Estimated insert size: 231000; agarose-fp estimation  
Estimated insert size: 315308; sum-of-contigs estimation  
Quality coverage: 6.76 in Q20 bases; agarose-fp estimation  
Quality coverage: 4.95 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 65 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1059: contig of 1059 bp in length  
\* 1060 1159: gap of unknown length  
\* 1160 2435: contig of 1276 bp in length  
\* 2436 2535: gap of unknown length  
\* 2536 3565: contig of 1030 bp in length  
\* 3566 3665: gap of unknown length  
\* 3666 4694: contig of 1029 bp in length  
\* 4695 4794: gap of unknown length  
\* 4795 6024: contig of 1230 bp in length  
\* 6025 6124: gap of unknown length  
\* 6125 7295: contig of 1171 bp in length  
\* 7296 7395: gap of unknown length  
\* 7396 8587: contig of 1192 bp in length  
\* 8588 8687: gap of unknown length  
\* 8688 10043: contig of 1356 bp in length  
\* 10044 10143: gap of unknown length  
\* 10144 11453: contig of 1310 bp in length  
\* 11454 11553: gap of unknown length  
\* 11554 12767: contig of 1214 bp in length  
\* 12768 12867: gap of unknown length  
\* 12868 14070: contig of 1203 bp in length  
\* 14071 14170: gap of unknown length  
\* 14171 15238: contig of 1068 bp in length  
\* 15239 15338: gap of unknown length  
\* 15339 16901: contig of 1563 bp in length  
\* 16902 17001: gap of unknown length  
\* 17002 18477: contig of 1476 bp in length  
\* 18478 18577: gap of unknown length  
\* 18578 20010: contig of 1433 bp in length  
\* 20011 20110: gap of unknown length  
\* 20111 21739: contig of 1629 bp in length  
\* 21740 21840: gap of unknown length  
\* 21840 24230: contig of 2391 bp in length  
\* 24231 24330: gap of unknown length  
\* 24331 26064: contig of 1734 bp in length  
\* 26065 26164: gap of unknown length  
\* 26165 27278: contig of 1114 bp in length  
\* 27279 27378: gap of unknown length  
\* 27379 29439: contig of 2061 bp in length  
\* 29440 29539: gap of unknown length  
\* 29540 32551: contig of 2812 bp in length  
\* 32552 32451: gap of unknown length  
\* 32452 34348: contig of 1897 bp in length  
\* 34349 34448: gap of unknown length  
\* 34449 36449: contig of 2001 bp in length  
\* 36450 36549: gap of unknown length  
\* 36550 37925: contig of 1376 bp in length  
\* 37926 38025: gap of unknown length  
\* 38026 39788: contig of 1763 bp in length  
\* 39789 39888: gap of unknown length  
\* 39889 42604: contig of 2716 bp in length  
\* 42605 42704: gap of unknown length  
\* 42705 44712: contig of 2008 bp in length  
\* 44713 44812: gap of unknown length  
\* 44813 47566: contig of 2754 bp in length  
\* 47567 47666: gap of unknown length  
\* 47667 49373: contig of 1707 bp in length  
\* 49374 49473: gap of unknown length  
\* 49474 52688: contig of 3215 bp in length  
\* 52689 52788: gap of unknown length  
\* 52789 56081: contig of 3293 bp in length  
\* 56082 56181: gap of unknown length  
\* 56182 59381: contig of 3200 bp in length  
\* 59382 59481: gap of unknown length  
\* 59482 61906: contig of 2425 bp in length  
\* 61907 62006: gap of unknown length  
\* 62007 64488: contig of 2482 bp in length  
\* 64489 64588: gap of unknown length  
\* 64589 65829: contig of 1241 bp in length  
\* 65830 65929: gap of unknown length  
\* 65930 68055: contig of 2126 bp in length  
\* 68056 68155: gap of unknown length  
\* 68156 72511: contig of 4356 bp in length  
\* 72512 72611: gap of unknown length  
\* 72612 76828: contig of 4217 bp in length  
\* 76829 76928: gap of unknown length  
\* 76929 80537: contig of 3609 bp in length  
\* 80538 80637: gap of unknown length  
\* 80639 86510: contig of 5873 bp in length  
\* 86511 86610: gap of unknown length  
\* 86611 88767: contig of 2157 bp in length  
\* 88768 88867: gap of unknown length  
\* 88868 91586: contig of 2719 bp in length  
\* 91587 91686: gap of unknown length  
\* 91687 96792: contig of 5106 bp in length  
\* 96793 96892: gap of unknown length  
\* 96893 102331: contig of 5339 bp in length  
\* 102332 102331: gap of unknown length  
\* 102332 106317: contig of 3986 bp in length  
\* 106318 106417: gap of unknown length  
\* 106419 111040: contig of 4623 bp in length  
\* 111041 111140: gap of unknown length  
\* 111141 116735: contig of 5595 bp in length  
\* 116736 116835: gap of unknown length  
\* 116836 120801: contig of 3966 bp in length  
\* 120802 120901: gap of unknown length  
\* 120902 127537: contig of 6636 bp in length  
\* 127538 127637: gap of unknown length  
\* 127638 133314: contig of 5677 bp in length  
\* 133315 133414: gap of unknown length  
\* 133415 139072: contig of 5658 bp in length  
\* 139073 139172: gap of unknown length  
\* 139173 143714: contig of 4542 bp in length  
\* 143715 143814: gap of unknown length  
\* 143815 151343: contig of 7529 bp in length  
\* 151344 151443: gap of unknown length  
\* 151443 160287: contig of 8744 bp in length  
\* 160288 160888: contig of 8401 bp in length  
\* 160889 168788: gap of unknown length

168789 177395: contig of 8607 bp in length  
\* 177396 177495: gap of unknown length  
\* 177496 189516: contig of 12021 bp in length  
\* 189517 189616: gap of unknown length  
\* 189617 204042: contig of 14426 bp in length  
\* 204043 204142: gap of unknown length  
\* 204143 214352: contig of 10210 bp in length  
\* 214353 214452: gap of unknown length  
\* 214453 225667: contig of 11215 bp in length  
\* 225668 225767: gap of unknown length  
\* 225768 245517: contig of 19750 bp in length  
\* 245518 259483: contig of 13866 bp in length  
\* 259484 259583: gap of unknown length  
\* 259584 272038: contig of 12456 bp in length  
\* 272039 272139: gap of unknown length  
\* 272140 292269: contig of 20130 bp in length  
\* 292270 321708: contig of 29339 bp in length.  
\* 292370 321708: contig of 29339 bp in length.

FEATURES  
source  
1. .321708  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-21119"  
/clone\_11b="RP23 mouse BAC library 23"

ORIGIN  
Query Match 1.4%; Score 37; DB 2; Length 321708;  
Best Local Similarity 100.0%; Pred. No. 2e-08;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 TTCCCGAGCAATTAAGTGTCTCTGAGATCACTGTA 274  
|||||  
DB 202208 TTCCCGAGCAATTAAGTGTCTCTGAGATCACTGTA 202244  
|||||

RESULT 33  
AC087159 338116 bp DNA linear HTG 09-DEC-2000  
LOCUS AC087159  
DEFINITION Mus musculus clone RP23-69F2, WORKING DRAFT SEQUENCE, 56 unordered  
pieces.  
AC087159.1 GI:11610885  
VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 338116)  
DOE Joint Genome Institute.  
2 (bases 1 to 338116)  
DOE Joint Genome Institute.  
Direct Submission.  
Submitted (09-DEC-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1763490  
Center clone name: RP23-69F2  
-----  
Summary Statistics  
Consensus quality: 297593 bases at least Q40  
Consensus quality: 310387 bases at least Q30  
Consensus quality: 317713 bases at least Q20  
Estimated insert size: 24100; agarose-fp estimation  
Estimated insert size: 332616; sum-of-contigs estimation

Quality coverage: 8.79 in Q20 bases; agarose-fp estimation  
Quality coverage: 6.37 in Q20 bases; sum-of-contigs estimation.  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 56 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1568: contig of 1568 bp in length  
\* 1569 1669: gap of unknown length  
\* 1669 2704: contig of 1036 bp in length  
\* 2705 2804: gap of unknown length  
\* 2805 4009: contig of 1205 bp in length  
\* 4010 4109: gap of unknown length  
\* 4110 5400: contig of 1291 bp in length  
\* 5401 5500: gap of unknown length  
\* 5501 6794: contig of 1293 bp in length  
\* 6794 6893: gap of unknown length  
\* 6894 8357: contig of 1464 bp in length  
\* 8358 8457: gap of unknown length  
\* 8458 9961: contig of 1503 bp in length  
\* 9961 10060: gap of unknown length  
\* 10061 11179: contig of 1119 bp in length  
\* 11180 11279: gap of unknown length  
\* 11280 12550: contig of 1271 bp in length  
\* 12551 12650: gap of unknown length  
\* 12651 14127: contig of 1477 bp in length  
\* 14128 14227: gap of unknown length  
\* 14228 15366: contig of 1139 bp in length  
\* 15367 15466: gap of unknown length  
\* 15467 16821: contig of 1355 bp in length  
\* 16822 16921: gap of unknown length  
\* 16922 18499: contig of 1578 bp in length  
\* 18500 18599: gap of unknown length  
\* 18600 19737: contig of 1138 bp in length  
\* 19738 19837: gap of unknown length  
\* 19838 22127: contig of 2230 bp in length  
\* 22128 22227: gap of unknown length  
\* 22228 24753: contig of 2526 bp in length  
\* 24754 24853: gap of unknown length  
\* 24854 26957: contig of 2104 bp in length  
\* 26958 27057: gap of unknown length  
\* 27058 30024: contig of 2967 bp in length  
\* 30025 30124: gap of unknown length  
\* 30125 32639: contig of 2515 bp in length  
\* 32640 32739: gap of unknown length  
\* 32740 34710: contig of 1971 bp in length  
\* 34711 34810: gap of unknown length  
\* 34811 38720: contig of 3910 bp in length  
\* 38721 38820: gap of unknown length  
\* 38821 42085: contig of 3265 bp in length  
\* 42086 42185: gap of unknown length  
\* 42186 45949: contig of 3764 bp in length  
\* 45950 46049: gap of unknown length  
\* 46050 49056: contig of 3007 bp in length  
\* 49057 49156: gap of unknown length  
\* 49157 51327: contig of 2171 bp in length  
\* 51328 51427: gap of unknown length  
\* 51428 54924: contig of 3497 bp in length  
\* 54925 55024: gap of unknown length  
\* 55025 57807: contig of 2783 bp in length  
\* 57808 57907: gap of unknown length  
\* 57909 59837: contig of 1930 bp in length  
\* 59838 59937: gap of unknown length  
\* 59939 63506: contig of 3569 bp in length  
\* 63507 63606: gap of unknown length  
\* 63607 65893: contig of 2287 bp in length  
\* 65894 65993: gap of unknown length  
\* 65994 68320: contig of 2327 bp in length  
\* 68321 70667: gap of unknown length  
\* 70667 70667: contig of 2447 bp in length

```

* 70968 70967: gap of unknown length
* 70968 73224: contig of 2257 bp in length
* 73325 73324: gap of unknown length
* 73325 79062: contig of 5738 bp in length
* 79063 79162: gap of unknown length
* 79163 85875: contig of 6713 bp in length
* 85876 85975: gap of unknown length
* 85976 93569: contig of 7594 bp in length
* 93570 93679: gap of unknown length
* 93679 100230: contig of 6561 bp in length
* 100231 100330: gap of unknown length
* 100331 110643: contig of 10313 bp in length
* 110644 110743: gap of unknown length
* 110744 119147: contig of 8404 bp in length
* 119148 119247: gap of unknown length
* 119248 128090: contig of 8843 bp in length
* 128091 128190: gap of unknown length
* 128191 134380: contig of 6190 bp in length
* 134381 134480: gap of unknown length
* 134481 142078: contig of 7598 bp in length
* 142079 142178: gap of unknown length
* 142179 152214: contig of 10036 bp in length
* 152215 152314: gap of unknown length
* 152315 158336: contig of 6322 bp in length
* 158337 158736: gap of unknown length
* 158737 167288: contig of 8552 bp in length
* 167289 167388: gap of unknown length
* 167389 177485: contig of 10097 bp in length
* 177486 177585: gap of unknown length
* 177586 191370: contig of 13785 bp in length
* 191371 191470: gap of unknown length
* 191471 204484: contig of 13014 bp in length
* 204485 209828: gap of unknown length
* 209829 209928: gap of unknown length
* 209929 224035: contig of 14107 bp in length
* 224036 224135: gap of unknown length
* 224136 238966: contig of 14831 bp in length
* 238967 239066: gap of unknown length
* 239067 252644: contig of 16198 bp in length
* 252645 255364: gap of unknown length
* 255365 257695: contig of 12331 bp in length
* 257696 267795: gap of unknown length
* 267796 286490: contig of 18695 bp in length
* 286491 307496: gap of unknown length
* 307497 307596: gap of unknown length
* 307597 338116: contig of 30520 bp in length.

```

```

FEATURES
Source
1.338116
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-69F2"
/clone_1b="RPI mouse BAC library 23"

```

## ORIGIN

```

Query Match      1.4%; Score 37; DB 2; Length 338116;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      238 TTCCCGACAAATAAGTGTCTCTGAGATCACTGTA 274
Db      189868 TTCCCGACAAATAAGTGTCTCTGAGATCACTGTA 189904

```

```

RESULT 34
AC125898/c 202792 bp DNA linear HTG 15-NOV-2002
LOCUS      Rattus norvegicus clone CH230-314P18, *** SEQUENCING IN PROGRESS
DEFINITION
AC125898
AC125898
AC125898.3 GI:25008623

```

```

KEYWORDS
SOURCE      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
ORGANISM    Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 202792)
Munzly,D.Marie, Metker,M.Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alebrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Bernamed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,T., Ceaar,H., Ceter,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,D.,
Dayla,M.L., Davis,C., Davy-Carroll,L., De Ande,C., Dederich,D.,
Delgado,O., Denon,S., Derramo,C., Ding,Y., Dinh,H., Divya,K.,
Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Guaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louised,H., Lorado,R.J., Lu,X., Ma,J.,
Machshawi,M., Mahidartine,M., Mamoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackelembi,O., Okwomu,G., Olarunpaagun,A., Pal,S., Parke,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Polndexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quito,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Ruiz,S.J.,
Rivers,C., Rodery,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Shen,H.,
Shedder,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smjs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Woodem,H., Wolley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von
Weinhausen,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23807727.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas

```

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

#### ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GWMO

Center clone name: CH230-314P18

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 166637 bases at least Q40

Consensus quality: 168395 bases at least Q30

Consensus quality: 169698 bases at least Q20

Estimated insert size: 172322; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/Genbankdraftdata.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 163896: contig of 163896 bp in length  
 \* 163897 163995: gap of unknown length  
 \* 16397 170075: contig of 6079 bp in length  
 \* 170076 170175: gap of unknown length  
 \* 170176 173847: contig of 3672 bp in length  
 \* 173848 173947: gap of unknown length  
 \* 173948 177191: contig of 3244 bp in length  
 \* 177192 177291: gap of unknown length  
 \* 177292 178689: contig of 1398 bp in length  
 \* 178690 178789: gap of unknown length  
 \* 178790 180421: contig of 1632 bp in length  
 \* 180422 180521: gap of unknown length  
 \* 180522 202792: contig of 22271 bp in length.

#### FEATURES

##### source

1. .202792  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-314P18"  
 29992. .31901  
 /note="wgs\_end\_extension  
 clone\_end:T7"  
 104416. .104486  
 /note="clone\_boundary  
 clone\_end:T7"  
 site:  
 end\_sequence:BZ249453"  
 122630. .127621  
 /note="clone boundary  
 clone\_end:Sp6  
 site:  
 end\_sequence:BZ249454"  
 163997. .165926  
 /note="wgs\_end\_extension  
 clone\_end:Sp6"  
 167885. .170075  
 /note="wgs\_end\_extension

##### ORIGIN

###### Query Match

Best Local Similarity 100.0%; Pred. No. 1e-06;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

##### QY

241 CCCAGCAATTAAGTGTCTCTGAGATCACTGTA 274  
 86522 CCCAGCAATTAAGTGTCTCTGAGATCACTGTA 86489

##### RESULT 35

###### LOCUS

BV002855 601 bp DNA linear STS 07-SEP-2002  
 S208B639RP11.T0 129S1/SVIMJ Mus musculus STS genomic, sequence  
 tagged site.

###### ACCESSION

BV002855  
 BV002855.1 GI:22755932

###### VERSIONS

KEYWORDS  
 STS  
 Mus musculus (house mouse)

###### SOURCE

Mus musculus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

###### REFERENCE

MADE, C.  
 Polymorphism Structure in the Mouse  
 Unpublished (2002)

###### TITLE

Contact: Kerstin Lindblad-Toh  
 Whitehead Institute for Biomedical Research, Center for Genome  
 Research  
 320 Charles Street, Cambridge, MA 02141, USA  
 Tel: 6172521477  
 Fax: 6172580903  
 Email: kersti@genome.wi.mit.edu

###### JOURNAL

Primer A: None  
 Primer B: None  
 STS size: 601  
 Protocol:

###### COMMENT

WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
 were generated from 129S1/SVIMJ, C3H/HeJ, and BALB/cByJ. The WGS  
 reads were placed uniquely on the MGS/CV3 C57BL/6J assembly and SNP  
 detection was carried out by SSANA-SNP. 225,000 reads were  
 annotated  
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J  
 and the strain from which the particular read came. The validation  
 rate for these SNPs was estimated at approximately 98%.

##### FEATURES

##### source

1. .601  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129S1/SVIMJ"  
 /db\_xref="taxon:10090"  
 /map="5 22-555 107871915-107872448"  
 /clone\_11b="129S1/SVIMJ"  
 <1. .>601  
 Location/Qualifiers

##### ORIGIN

###### STS

Query Match 1.2%; Score 32; DB 11; Length 601;  
 Best. Local Similarity 100.0%; Pred. No. 7.1e-06;





## FEATURES

```

* 28550 33402: contig of 4853 bp in length
* 33403 33502: gap of unknown length
* 33503 33568: contig of 4066 bp in length
* 33569 37668: gap of unknown length
* 42544 42544: contig of 4876 bp in length
* 42645 42645: gap of unknown length
* 47232 47232: contig of 4587 bp in length
* 47332 52835: gap of 5504 bp in length
* 52836 52935: gap of unknown length
* 52936 59373: contig of 6438 bp in length
* 59374 59473: gap of unknown length
* 59474 66585: contig of 7112 bp in length
* 66586 66685: gap of unknown length
* 66686 74608: contig of 7923 bp in length
* 74609 74708: gap of unknown length
* 74709 75327: contig of 619 bp in length
* 75328 75427: gap of unknown length
* 75428 76074: contig of 647 bp in length
* 76075 76174: gap of unknown length
* 76175 81491: contig of 5317 bp in length
* 81492 81591: gap of unknown length
* 81592 88018: contig of 6427 bp in length
* 88019 88118: gap of unknown length
* 88119 98670: contig of 10552 bp in length
* 98671 98770: gap of unknown length
* 98771 113277: contig of 14507 bp in length
* 113278 113377: gap of unknown length
* 113378 127626: contig of 14249 bp in length
* 127627 127727: gap of unknown length
* 127727 146516: contig of 18790 bp in length.
Location/Qualifiers
1. 146516
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-21506"
1. 1169
/feature="assembly_name:Contig19"
1270. 2651
/feature="assembly_name:Contig21"
2752. 4398
/feature="assembly_name:Contig22"
4499. 5938
/feature="assembly_name:Contig23"
6039. 8062
/feature="assembly_name:Contig24"
8163. 10520
/feature="assembly_name:Contig25"
10621. 13358
/feature="assembly_name:Contig26"
13459. 15803
/feature="assembly_name:Contig27"
15904. 18487
/feature="assembly_name:Contig28"
18588. 21397
/feature="assembly_name:Contig29"
21498. 25904
/feature="assembly_name:Contig30"
26005. 28449
/feature="assembly_name:Contig31"
28550. 33402
/feature="assembly_name:Contig32"
33503. 37568
/feature="assembly_name:Contig33"
37669. 42544
/feature="assembly_name:Contig34"
42645. 47231
/feature="assembly_name:Contig35"
47332. 52835
/feature="assembly_name:Contig36"
52936. 59373

```

## ORIGIN

```

Query Match 1.0%; Score 28; DB 2; Length 146516;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

misc_feature /note="assembly_name:Contig37"
59474..66585
/feature="assembly_name:Contig38"
66686..74608
misc_feature /note="assembly_name:Contig39"
74709..75327
misc_feature /note="assembly_name:Contig4"
75428..76074
/feature="assembly_name:Contig3"
76175..81491
misc_feature /note="assembly_name:Contig40"
81592..88018
/feature="assembly_name:Contig41"
88119..98670
misc_feature /note="assembly_name:Contig42"
98771..113277
misc_feature /note="assembly_name:Contig43"
113378..127626
misc_feature /note="assembly_name:Contig44"
127727..146516
/feature="assembly_name:Contig45"

```

```

Qy 2652 ACBAATTGACAAAAAAAAAAAAAAAAAAAA 2679
Db 102453 ACBAATTGACAAAAAAAAAAAAAAAAAAAA 102426

```

## RESULT 38

```

AC126634 218226 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-13813, WORKING DRAFT SEQUENCE.
DEFINITION AC126634
ACCESSION AC126634
VERSION AC126634.4 GI:25007898
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 218226)
REFERENCE
AUTHORS Muzny,D,Marie, Metzker,M, Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amth, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crease, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Deigado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Diya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Hayes, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunnarsson, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowals, C., Kratz, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Louieged, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

```



Shetty, J., Shvaritsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Szatek, A., Tabore, P., Taylor, C.,  
Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umali, K.,  
Vals, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J.,  
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., Wiley, F.,  
Williams, G., Wilson, R., Wlezyk, R., Wooden, H., Wortley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Wernstock, G., and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 228987)  
Worley, K.C.

Direct Submission  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 228987)

Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265947.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
<http://www.hgsc.bcm.tmc.edu/projects/rat/>. Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.003; Length 228987;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6876 ACAATTGCACCAAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
|||||

RESULT 40  
AC097665/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

1 (bases 1 to 274796)  
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Aisbrouck, S., Amin, A., Angiano, D., Ayala-Bechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bernal, K., Blair, J., Blankensbury, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy, Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Gasta, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huyl, S., Hume, J., Idelbr, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebwohl, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzuela, L., Louieged, H., Lozadou, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenan, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervill, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokwelen, O., Okwuonu, G., Olamunungu, A., Pal, S., Parks, K., Pasternack, S., Paul, H., Perez, A., Perez, J., Pfamkoch, C.,

/note="wgs\_end\_extension  
clone\_end:T7"  
2887..3976  
/note="wgs\_end\_extension  
clone\_end:T7"  
complement (4263..4996)  
/note="clone\_boundary  
clone\_end:T7  
site:  
end\_sequence: BH258473"  
complement (227308..228112)  
/note="clone\_boundary  
clone\_end:Sp6  
site:  
end\_sequence: BH258474"

AC097665 274796 bp DNA linear HTG 10-MAY-2003  
Rattus norvegicus clone CH230-86C24, WORKING DRAFT SEQUENCE, 4  
unordered pieces.  
AC097665  
AC097665.7 GI:30520736  
HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Rivers, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivas, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Riz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, D., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Vales, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wiczek, R., Woden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, D., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 274796)  
Worley, K.C.

Direct Submission  
Submitted (21-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 274796)  
Rat Genome Sequencing Consortium.

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:25009660. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: GPUL  
Center clone name: CH230-86C24

----- Summary Statistics  
Assembly program: Atlas 3.0  
Consensus quality: 262493 bases at least Q40  
Consensus quality: 263877 bases at least Q30  
Consensus quality: 264932 bases at least Q20  
Estimated insert size: 270537; sum-of-contigs estimation  
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
NOTE: This sequence may represent more than one clone.  
NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 271424: contig of 271424 bp in length  
\* 271425 271524: gap of unknown length  
\* 271525 272554: contig of 1030 bp in length

272555 272654: gap of unknown length  
\* 272655 273674: contig of 1020 bp in length  
\* 273675 273774: gap of unknown length  
\* 273775 274796: contig of 1022 bp in length.

FEATURES  
source  
1..274796  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-86C24"  
1..1265  
/note="wgs\_contig"

ORIGIN  
misc\_feature

Query Match  
Best Local Similarity 100.0%; Score 28; DB 2; Length 274796;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2652 ACMAATTGCACMAAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
Db 75393 ACMAATTGCACMAAAAAAAAAAAAAAAAAAAAA 75366

RESULT 41  
AC117568 167166 bp DNA linear HTG 27-AUG-2002  
LOCUS  
AC117568  
DEFINITION Mus musculus clone RP24-511J14, WORKING DRAFT SEQUENCE, 4 unordered pieces.  
AC117568.3 GI:22507215  
VERSION  
HTG, HTGS\_PHASE1, HTGS\_DRAFT, HTGS\_FULFILL.  
KEYWORDS  
Mus musculus (house mouse)  
SOURCE  
ORIGINISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 (bases 1 to 167166)  
Birren, B., Nusbaum, C. and Lander, E.  
Mus musculus, clone RP24-511J14  
Unpublished  
2 (bases 1 to 167166)

JOURNAL  
TITLE  
AUTHORS  
REFERENCES

2 (bases 1 to 167166)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Bonkhgalter, B., Brown, A., Camarata, J., Campione, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, U.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Lacroque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rie, C., Rogov, P., Roman, N., Roselt, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 167166)

REFERENCE  
JOURNAL  
TITLE  
AUTHORS  
REFERENCES

Birren, B., Nusbaum, C., Lander, E., Allen, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Bonkhgalter, B., Brown, A., Camarata, J., Campione, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, U.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (27-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 27, 2002 this sequence version replaced gi:22296976.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L21336

Center clone name: 511.J.14

## Summary Statistics

Sequencing vector: Plasmid, n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 165942 bases at least Q40

Consensus quality: 166212 bases at least Q30

Consensus quality: 166486 bases at least Q20

Insert size: 163000; agarose-fp

Insert size: 166866; sum-of-coverage

Quality coverage: 12.9 in Q20 bases; agarose-fp

Quality coverage: 12.6 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 18065: contig of 18065 bp in length  
\* 18066 18165: gap of 100 bp  
\* 18166 24868: contig of 6703 bp in length  
\* 24869 24968: gap of 100 bp  
\* 24969 71883: contig of 46915 bp in length  
\* 71884 71983: gap of 100 bp  
\* 71984 167166: contig of 95183 bp in length.

## FEATURES

## source

1. 167166

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/clone="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

/clone\_lib="RP24-511J14"

Query Match 1.0%; Score 27; DB 2; Length 167166;

Best Local Similarity 100.0%; Pred. No. 0.013; Indels 0; Gaps 0;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2653 CAAATTGACAAAAA 2679

60190 CAAATTGACAAAAA 60216

60190 CAAATTGACAAAAA 60216

## RESULT 42

## AC115904

## LOCUS

## DEFINITION

## AC115904

## AC115904.13

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 167938)

Birren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bachtin, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeRellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagoos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Scudis, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission







```

repeat_region      /rpt_family="ID"
26042..26434      /rpt_family="MaLR"
repeat_region      /rpt_family="L1"
26449..26560      /rpt_family="L1"
repeat_region      /rpt_family="B2"
26772..26968      /rpt_family="B2"
repeat_region      /rpt_family="Alu"
26981..27103      /rpt_family="Alu"
repeat_region      /rpt_family="MaLR"
27126..27238      /rpt_family="MaLR"
repeat_region      /rpt_family="Alu"
27144..27276      /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
27892..28269      /rpt_family="MaLR"
repeat_region      /rpt_family="MaLR"
28623..28744      /rpt_family="B2"
repeat_region      /rpt_family="B2"
29031..29137      /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
29414..30117      /rpt_family="MER1_type"
repeat_region      /rpt_family="Alu"
30921..31047      /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
37477..37709      /rpt_family="B4"
repeat_region      /rpt_family="B4"
37710..37869      /rpt_family="B4"
repeat_region      /rpt_family="B4"
37870..37913      /rpt_family="B4"
repeat_region      /rpt_family="B4"
38345..38838      /rpt_family="MaLR"
repeat_region      /rpt_family="MaLR"
40401..40617      /rpt_family="ERV1"
repeat_region      /rpt_family="ERV1"
40733..40863      /rpt_family="MaLR"
repeat_region      /rpt_family="MaLR"
40864..41239      /rpt_family="ERV1"
repeat_region      /rpt_family="ERV1"
41480..41541      /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
41540..41685      /rpt_family="Alu"
repeat_region      /rpt_family="Alu"
43155..43553      /rpt_family="MaLR"
repeat_region      /rpt_family="MaLR"
46290..46389      /rpt_family="L2"
repeat_region      /rpt_family="L2"
48112..48613      /rpt_family="MaLR"
repeat_region      /rpt_family="MaLR"
49627..50013      /rpt_family="MaLR"
repeat_region      /rpt_family="MaLR"
50710..50882      /rpt_family="MER2_type"
repeat_region      /rpt_family="MER2_type"
50890..51030      /rpt_family="MER2_type"
repeat_region      /rpt_family="MER2_type"
51031..51192      /rpt_family="Achoho"
repeat_region      /rpt_family="Achoho"
53925..54042      /rpt_family="MER1_type"
repeat_region      /rpt_family="MER1_type"
54477..54561      /rpt_family="L1"

```

Query Match 1.0%; Score 27; DB 10; Length 221753;  
 Best Local Similarity 100.0%; Pred. No. 0.012; Mismatches 0; Indels 0; Gaps 0;

2653 CAAATTGACAAAAA 2679  
 DB 49641 CAAATTGACAAAAA 49615

RESULT 44 AC123851 230221 bp DNA linear ROD 11-NOV-2003  
 LOCUS AC123851/c AC123851  
 DEFINITION Mus musculus BAC clone RP23-323K4 from chromosome 16, complete

```

ACCESSION      AC123851
VERSION        AC123851.4
KEYWORDS       GI:28376830
SOURCE         HTG.
ORGANISM       Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 230221)
AUTHORS        Shah,N., Kozlowski,A. and Schatzkammer,K.
TITLE          The sequence of Mus musculus BAC clone RP23-323K4
JOURNAL        Unpublished (2001)
REFERENCE      2 (bases 1 to 230221)
AUTHORS        Wilson,R.
TITLE          Sequencing of Mus musculus
JOURNAL        Unpublished (2001)
REFERENCE      3 (bases 1 to 230221)
AUTHORS        McPherson,J.D. and Waterson,R.H.
TITLE          Direct Submission
JOURNAL        Submitted (01-JUN-2002) Genome Sequencing Center, 4444 Forest Park
                Parkway, St. Louis, MO 63108, USA
REFERENCE      4 (bases 1 to 230221)
AUTHORS        McPherson,J.D. and Waterson,R.H.
TITLE          Direct Submission
JOURNAL        Submitted (14-FEB-2003) Genome Sequencing Center, 4444 Forest Park
                Parkway, St. Louis, MO 63108, USA
REFERENCE      5 (bases 1 to 230221)
AUTHORS        Wilson,R.
TITLE          Direct Submission
JOURNAL        Submitted (11-NOV-2003) Department of Genetics, Washington
                University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
                On Feb 14, 2003 this sequence version replaced gi:23238091.
COMMENT        ----- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu
                Contact: submissions@watson.wustl.edu
                ----- Summary Statistics
                Center project name: M_BA0323K04

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:  
 The RP23-323 BAC library has been constructed by Kazuhiro Osegaawa and Minako Tatemio in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

# NEIGHBORING SEQUENCE INFORMATION: This sequence is the entire insert of the clone.

FEATURES  
source  
1. .230221  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="16"  
/map="16"  
/clone="RP23-323K4"  
/clone\_lib="RPCI-23"  
20. .154  
repeat\_region  
/rpt\_family="B2"  
182. .339  
repeat\_region  
/rpt\_family="B2"  
363. .431  
repeat\_region  
/rpt\_family="tRNA-Ala-GCY"  
432. .452  
repeat\_region  
/rpt\_family="BCI\_MM"  
1466. .1578  
repeat\_region  
/rpt\_family="B4"  
1900. .2161  
repeat\_region  
/rpt\_family="MALR"  
2730. .2862  
repeat\_region  
/rpt\_family="B4"  
2881. .2975  
repeat\_region  
/rpt\_family="B2"  
2976. .3162  
repeat\_region  
/rpt\_family="B2"  
3223. .3324  
repeat\_region  
/rpt\_family="MIR1\_cype"  
3376. .3473  
repeat\_region  
/rpt\_family="Alu"  
3589. .3802  
repeat\_region  
/rpt\_family="MALR"  
3805. .4093  
repeat\_region  
/rpt\_family="B4"  
4850. .5027  
repeat\_region  
/rpt\_family="B2"  
5068. .5125  
repeat\_region  
/rpt\_family="MIR"  
5138. .5177  
repeat\_region  
/rpt\_family="ID"  
5306. .5460  
repeat\_region  
/rpt\_family="MALR"  
5486. .5756  
repeat\_region  
/rpt\_family="MALR"  
5757. .5797  
repeat\_region  
/rpt\_family="B2"  
5947. .6185  
repeat\_region  
/rpt\_family="B4"  
6193. .6274  
repeat\_region  
/rpt\_family="B4"  
6275. .6421  
repeat\_region  
/rpt\_family="Alu"  
6422. .6564  
repeat\_region  
/rpt\_family="B4"  
6835. .6908  
repeat\_region  
/rpt\_family="Alu"  
6976. .7122  
repeat\_region  
/rpt\_family="B4"  
7123. .7360  
repeat\_region  
/rpt\_family="B4"  
7361. .7431  
repeat\_region  
/rpt\_family="B4"  
8098. .8198  
repeat\_region  
/rpt\_family="B4"  
8755. .8897  
repeat\_region  
/rpt\_family="Alu"  
9653. .9679  
repeat\_region  
/rpt\_family="B4"  
9680. .9812  
repeat\_region  
/rpt\_family="Alu"

repeat\_region 10174. .10219  
/rpt\_family="Alu"  
repeat\_region 10220. .10389  
/rpt\_family="MALR"  
repeat\_region 10590. .10652  
/rpt\_family="Alu"  
repeat\_region 10695. .10819  
/rpt\_family="B4"  
repeat\_region 11328. .11502  
/rpt\_family="L1"  
repeat\_region 11589. .11645  
/rpt\_family="B4"  
repeat\_region 11643. .11699  
/rpt\_family="B4"  
repeat\_region 11700. .11830  
/rpt\_family="Alu"  
repeat\_region 12293. .12395  
/rpt\_family="Alu"  
repeat\_region 12558. .12755  
/rpt\_family="B2"  
repeat\_region 13117. .13341  
/rpt\_family="B4"  
repeat\_region 13328. .13461  
/rpt\_family="B4"  
repeat\_region 13350. .13469  
/rpt\_family="Alu"  
repeat\_region 13506. .13579  
/rpt\_family="B4"  
repeat\_region 13942. .13992  
/rpt\_family="MIR"  
repeat\_region 15276. .15865  
/rpt\_family="MALR"  
repeat\_region 16313. .16451  
/rpt\_family="Alu"  
repeat\_region 16687. .16891  
/rpt\_family="B2"  
repeat\_region 16892. .16945  
/rpt\_family="tRNA-Ala-GCY"  
repeat\_region 17061. .17359  
/rpt\_family="L1"  
repeat\_region 17361. .17554  
/rpt\_family="L1"  
repeat\_region 17578. .17996  
/rpt\_family="ERVK"  
repeat\_region 17998. .18351  
/rpt\_family="ERVK"  
repeat\_region 18362. .19317  
/rpt\_family="ERVK"  
repeat\_region 19341. .20337  
/rpt\_family="ERVK"  
repeat\_region 20582. .21245  
/rpt\_family="ERVK"

Query Match 1.0%; Score 27; DB 10; Length 230221;  
Best Local Similarity 100.0%; Pred NO. 0.012; Mismatches 0; Indels 0; Gaps 0;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2653 CAAATTGACAAAAA 2679  
Db 221953 CAAATTGACAAAAA 221927  
RESULT 45  
AL772311/c 234393 bp DNA linear ROD 08-AUG-2003  
LOCUS Mouse DNA sequence from clone RP23-87P16 on chromosome 4, complete  
DEFINITION sequence.  
ACCESSION AL772311 GI:33412257  
VERSION AL772311.19  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 234393)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS Andrew.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-AUG-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 COMMENT On Aug 1, 2003 this sequence version replaced gi:33386469. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-67P16 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES  
 source  
 1.234393  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-67P16"  
 /clone\_1fb="RPCI-23"

ORIGIN

Query Match 1.0%; Score 27; DB 10; Length 234393;  
 Best Local Similarity 100.0%; Pwd.No. 0.012;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2653 CAAATTGACAAAAA 2679  
 |||||  
 Db 201565 CAAATTGACAAAAA 201559

RESULT 46  
 AC102564/c 239783 bp DNA linear HTG 10-Jul-2003  
 LOCUS Mus musculus clone RP23-210C12, WORKING DRAFT SEQUENCE, 6 unordered  
 DEFINITION Mus musculus clone RP23-210C12, WORKING DRAFT SEQUENCE, 6 unordered  
 AC102564  
 pieces.  
 AC102564  
 AC102564.3 GI:32490706  
 VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 239783)

AUTHORS Birren, B., Nusbaum, C. and Lander, E.  
 TITLE Mus musculus, clone RP23-210C12  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 239783)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Brown, A., Camarata, U., Campiolo, V., Boguslavsky, L., Boukhgalter, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dods, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, M., Illiv, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, D., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McKeethers, R., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Teafaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vaseliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
 JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 REFERENCE 3 (bases 1 to 239783)  
 AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelell, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardina, S., Graham, L., Grand-Pierre, N., Hafer, N., Hagopian, D., Hagos, B., Hall, V., Horton, L., Hulme, M., Illiv, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Rachupda, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Teafaye, S., Theodore, J., Topham, K., Travers, M., Vaseliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
 JOURNAL Submitted (10-Jul-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 COMMENT On Jul 10, 2003 this sequence version replaced gi:2281552. All repeats were identified using RepeatMasker:  
 Smit, A. P. A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WtBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 Project Information  
 Center project name: L19015  
 Center clone name: 210 C\_12  
 Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 238594 bases at least Q40  
 Consensus quality: 238993 bases at least Q30  
 Consensus quality: 239154 bases at least Q20  
 Insert size: 212000; agarose-fp

Insert size: 239283; sum-of-coverage  
Quality coverage: 10.7 in Q20 bases; agarose-fp  
Quality coverage: 9.5 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 5588: contig of 5588 bp in length  
5589 5689: gap of 100 bp  
5690 90331: contig of 84643 bp in length  
90332 90431: gap of 100 bp  
90432 120455: contig of 30024 bp in length  
120456 120555: gap of 100 bp  
120556 165384: contig of 44829 bp in length  
165385 165484: gap of 100 bp  
165485 228023: contig of 62339 bp in length  
228024 228123: gap of 100 bp  
228124 239783: contig of 11660 bp in length.

FEATURES  
source  
1. 239783  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-210C12"  
/clone\_1bp="RP23-210C12"  
1. 5588  
/note="assembly\_fragment"  
5689. 90331  
/note="assembly\_fragment"  
90432. 120455  
/note="assembly\_fragment"  
120556. 165384  
/note="assembly\_fragment"  
165485. 228023  
/note="assembly\_fragment"  
228124. 239783  
/note="assembly\_fragment"  
clone\_end:17  
vector\_side:right"

ORIGIN  
Query Match 1.0%; Score 27; DB 2; Length 239783;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2653 CAAATTGACCAAAAAAAAAAAAAA 2679  
DB 168024 CAAATTGACCAAAAAAAAAAAAAA 167998

RESULT 47  
AC129762 266632 bp DNA 1linear HTG 09-NOV-2002  
LOCUS Rattus norvegicus clone CH230-42A2, WORKING DRAFT SEQUENCE, 2  
DEFINITION  
AC129762  
unordered pieces.  
AC129762.3 GI:24635717  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 266632)  
REFERENCE  
AUTHORS Wuzny,D.,Marle., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooke,S., Amin,A., Angulano,D., Anyalebechi,V., Ayvagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benhamed,F.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (02-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 266632)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 6, 2002 this sequence version replaced gi:2326431.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Biswal,K., Blair,J., Blankenburg,K., Blych,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Cebasa,H., Center,A., Chacko,U., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinu,H., Diya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,I., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabiela,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Gillingham,M., Givara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenzulewa,L., Loubege,H., Lozada,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Mijca,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Muniasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwackelmehe,O., Okunolu,G., Olamunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Plankoch,C., Plopper,F., Polidexter,A., Popovic,D., Primus,E., Pu,L., L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,R., Shetty,D., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smales,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Soes,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanai,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wleczyk,R., Woodson,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G., and Gibbs,R.A.

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GFRN  
 Center clone name: CH230-42A2  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 231272 bases at least Q40  
 Consensus quality: 234148 bases at least Q30  
 Consensus quality: 235932 bases at least Q20  
 Estimated insert size: 240545; sum-of-coverage  
 Quality coverage: 7x in Q20 bases; sum-of-coverage estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

FEATURES  
 source  
 1. 266632  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-42A2"  
 1. 2475  
 /note="wgs\_end\_extension  
 clone\_end:T7"  
 2526..3711  
 /note="wgs\_end\_extension  
 clone\_end:T7"  
 9336..10170  
 /note="clone\_boundary  
 clone\_end:T7  
 site:BCORI  
 end\_sequence:BH345972"  
 48492..49582  
 /note="wgs\_contig"  
 219565..221299  
 /note="wgs contig"  
 complement(260236..260933)  
 /note="clone\_boundary  
 clone\_end:Sp6  
 site:BCORI  
 end\_sequence:BH345974"  
 263024..264609  
 /note="wgs\_end\_extension  
 clone\_end:Sp6"

ORIGIN  
 Query Match 1.0%; Score 27; DB 2; Length 266632;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2653 CAATTGACAAAAA 2679  
 |||||  
 DB 247436 CAATTGACAAAAA 247462  
 |||||

RESULT 48  
 AC115400/c 269117 bp DNA linear HTG 08-OCT-2002  
 LOCUS Rattus norvegicus clone CH230-59M6, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 DEFINITION 5 unordered pieces.

ACCESSION  
 AC115400  
 AC115400.4 GI:23270024  
 HTG; HTGS PHASE1; HTGS ENRICHED.  
 KEYWORDS  
 Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 269117)

REFERENCE  
 AUTHORS  
 Muzny D, Marie, Melker, M, Lee, Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Albrooke, S., Amin, A., Angiano, D.,  
 Anyalelechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
 Fernandez, S., Flanley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,  
 Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jolivet, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Karpach, S.,  
 Karpach, S., Kelly, S., Khan, Z., Kling, L., Kovar, C.,  
 Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Loshenshewa, L., Louissege, H., Lozano, R.J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martin, E.,  
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minje, E., Montemayor, J., Moore, S.,  
 Milosavljevic, A., Miner, G., Minje, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Mundana, N., Murphy, M., Natir, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwokeneme, O., Okwuonu, G., Olarunsgun, A., Pal, S., Parke, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Poldinger, A., Popovic, D., Prins, E., Pu, L.,  
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,  
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Slason, I., Sitter, C.D., Smajs, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorlie, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,  
 Valas, R., Vera, V., Villalana, D., Walidron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, X., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstein, G. and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 269117)  
 Worley, K.C.  
 Direct Submission  
 Submitted (19-MAR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, Houston, TX 77030, USA  
 3 (bases 1 to 269117)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (08-OCT-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Sep 23, 2002 this sequence version replaced gi:22450411.

REFERENCE  
 JOURNAL  
 TITLE  
 AUTHORS  
 COMMENT



The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequencing contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

#### Genome Center

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

#### Project Information

Center project name: GNMW  
Center clone name: CH230-59M6

#### Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 180274 bases at least Q40  
Consensus quality: 183104 bases at least Q30  
Consensus quality: 185338 bases at least Q20  
Estimated insert size: 198590; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
(see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).  
NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 86595: contig of 86595 bp in length  
\* 86596 86695: gap of unknown length  
\* 86696 167948: contig of 81253 bp in length  
\* 167949 168048: gap of unknown length  
\* 168049 265576: contig of 97528 bp in length  
\* 265577 265676: gap of unknown length  
\* 265677 266793: contig of 1117 bp in length  
\* 266794 266893: gap of unknown length  
\* 266894 269117: contig of 2224 bp in length.

#### FEATURES

##### source

1. 269117  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-59M6"  
1. 1682  
misc\_feature  
/note="wgs contig"  
95067..96319  
misc\_feature  
/note="wgs contig"  
125123..127013  
misc\_feature  
/note="wgs contig"

##### ORIGIN

Query Match 1.0%; Score 27; DB 2; Length 269117;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2653 CAATTGACAAAAAAAAAAAAAAAAA 2679  
|||||  
Db 42476 CAATTGACAAAAAAAAAAAAAAAA 42450  
|||||

RESULT 49  
AC131863 322833 bp DNA linear HTG 23-NOV-2002  
LOCUS AC131863 322833 bp DNA linear HTG 23-NOV-2002

#### DEFINITION

Rattus norvegicus clone CH230-433M17, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 10 unordered pieces.

#### ACCESSION

AC131863 GI:25188440

#### VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.

#### KEYWORDS

Rattus norvegicus

#### SOURCE

Rattus norvegicus

#### ORGANISM

Rattus norvegicus

#### REFERENCE

1 (bases 1 to 322833)

#### AUTHORS

Muzny,D,Maritz,M,Leese,Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benhammed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Casar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedrich,D, Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Evans,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Frazer,C,M, Gabiella,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Gridley,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,K, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogue,M, Hollins,B, Howells,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kratt,C,L, Labow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lounsbury,L, Loulsegod,H, Lozano,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Mallory,K, Mangum,A, Mangum,B, Mapa,P, Martin,K, Martin,R, Martinez,E, Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,B, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwackelmebe,O, Okunonu,G, Olumunagbon,A, Pal,S, Parks,K, Pasernak,S, Paul,H, Perez,A, Perez,L, Frankoch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shattman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajda,D, Sneed,A, Sodergren,B, Song,X,-Z, Sotelle,R, Sosa,D, Stemle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,D, Trejos,Z, Uman,I, K, Valas,R, Vera,V, Villalana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,K, Wleczek,R, Woodson,H, Woley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.

#### TITLE

Direct Submission

#### JOURNAL

Unpublished

#### REFERENCE

2 (bases 1 to 322833)

#### AUTHORS

Rat Genome Sequencing Consortium.

#### TITLE

Submitted (27-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

#### REFERENCE

3 (bases 1 to 322833)

#### AUTHORS

Rat Genome Sequencing Consortium.

#### TITLE

Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

## COMMENT

Baylor Plaza, Houston, TX 77030, USA  
 On Nov 23, 2002 this sequence version replaced gi:23322024.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

## ----- Genome Center

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center Project name: GVEC  
 Center clone name: CH230-433M17

## ----- Summary Statistics

Assembly program: Phrap; version 0.990329  
 Consensus quality: 272635 bases at least Q40  
 Consensus quality: 278236 bases at least Q30  
 Consensus quality: 281659 bases at least Q20  
 Estimated insert size: 275728; sum-of-contigs estimation  
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drafc\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drafc_data.html))  
 \* NOTE: This sequence may represent more than one clone.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 10 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 14872: contig of 14872 bp in length
* 14873 14972: gap of unknown length
* 14873 46078: gap of 31106 bp in length
* 46079 46178: gap of unknown length
* 46179 51691: contig of 5513 bp in length
* 51692 51791: gap of unknown length
* 51792 113737: contig of 61946 bp in length
* 113738 113837: gap of unknown length
* 113838 313242: contig of 199405 bp in length
* 313243 313243: gap of unknown length
* 313243 314356: contig of 1014 bp in length
* 314357 314456: gap of unknown length
* 314457 316001: contig of 1545 bp in length
* 316002 316101: gap of unknown length
* 316102 317157: contig of 1056 bp in length
* 317158 317257: gap of unknown length
* 317258 318770: contig of 1513 bp in length
* 318771 322833: contig of 3963 bp in length.
* 318871

```

## FEATURES

```

source      1..322833
             /organism="Rattus norvegicus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10116"
             /clone="CH230-433M17"
misc_feature 10649..11718
             /note="wgs contig"
misc_feature 13663..14872
             /note="wgs contig"
misc_feature 14973..16147
             /note="wgs contig"
misc_feature 21483..22546

```

## ORIGIN

```

misc_feature /note="wgs contig"
misc_feature 26003..29438
             /note="wgs contig"
misc_feature 39190..40851
             /note="wgs contig"
misc_feature 44314..46078
             /note="wgs contig"
misc_feature 46179..47549
             /note="wgs contig"
misc_feature 78005..79490
             /note="wgs contig"
misc_feature 82448..86781
             /note="wgs contig"
misc_feature 102318..103470
             /note="wgs contig"
misc_feature 103521..104683
             /note="wgs contig"
misc_feature 113838..114931
             /note="wgs contig"
misc_feature 119033..120040
             /note="wgs contig"
misc_feature 311394..313242
             /note="wgs contig"

```

## Query Match

Best Local Similarity 1.0%; Score 27; DB 2; Length 322833;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2653 CAATTGACAAAAA 2679

DB 170299 CAATTGACAAAAA 170325

## RESULT 50

ATH529544

## LOCUS

ATH529544 444 bp DNA linear PLN 29-MAR-2003

## DEFINITION

Arabidopsis thaliana T-DNA flanking sequence, left border, clone

## ACCESSION

AJ529544.1 GI:26797804

## VERSION

AJ529544 left border; T-DNA flanking sequence.

## KEYWORDS

Arabidopsis thaliana (thale cress)

## SOURCE

Arabidopsis thaliana

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## REFERENCE

1 Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,

## AUTHORS

Chavirin, S., Bechold, N., Cnuad, C., Derose, R., Pelletier, G.,

## TITLE

T-DNA integration into the Arabidopsis genome depends on sequences

## JOURNAL

EMBO Rep. 3 (12), 1152-1157 (2002)

## MEDLINE

22363535

## PUBMED

12446565

## REFERENCE

2 (bases 1 to 444)

## AUTHORS

Balzerque, S.

## TITLE

Direct Submision

## JOURNAL

Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue

## COMMENT

Gaston Cremieux, 91057 Evry cedex, FRANCE

## FEATURES

PCR was performed on DNA from transformants of Arabidopsis thaliana

## PLANTS

plants from INRA (Versailles). The DNA fragment(s) resulting from

## THE PCR

were directly sequenced from the left or the right border

## TO DETERMINE

the genomic sequence flanking the insertion. T-DNA

## DERIVED

sequences were removed. Information to order the

## CORRESPONDING

mutant line and a link to a database providing a

## GRAPHICAL

display of the insertion site are available at

## HTTP://DBSGAP

[versailles.inra.fr/publiclines/](http://dbsgap.versailles.inra.fr/publiclines/). This sequence has

## BEEN GENERATED

in the framework of the French plant genomics

## PROGRAM 'GENOPLANTE'

(<http://www.genoplante.com> and

## HTTP://GENOPLANTE

<http://genoplante-info.infobiogen.fr>).

## LOCATION/QUALIFIERS

Location/Qualifiers

## source

```
1..444
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Masilllewska"
/db_xref="taxon:3702"
/clone="189D04"
/clone_1b="Arabidopsis thaliana T-DNA insertion lines"
1..444
/note="T-DNA flanking sequence
left border"
```

## misc\_feature

```
1..444
/note="T-DNA flanking sequence
left border"
```

## ORIGIN

## Query Match

1.0%; Score 26; DB 8; Length 444;

Best Local Similarity 100.0%; Pred. No. 0.021; Mismatches 0; Indels 0; Gaps 0;

## Qy

2654 AAATTGACAAAAAAAAAAAAAAAAAAAA 2679

## Db

390 AAATTGACAAAAAAAAAAAAAAAAAAAA 415

Search completed: January 15, 2005, 10:12:12  
Job time : 11210 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2005, 00:57:51 : Search time 1230 Seconds

(without alignments)  
11433.495 Million cell updates/sec

Title: US-10-048-046-1

Perfect score: 2679  
Sequence: 1 aagaatcgcgcagcagccg.....acaaaaaaaaaaaaaaaaa 2679

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 413486 seqs, 2624710521 residues

Word size : 20

Total number of hits satisfying chosen parameters: 18399

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

N\_Geneseq\_23Sep04:\*  
1: genebegn1980s:\*  
2: genebegn1980s:\*  
3: genebegn2000s:\*  
4: genebegn2001as:\*  
5: genebegn2001as:\*  
6: genebegn2002as:\*  
7: genebegn2002as:\*  
8: genebegn2003as:\*  
9: genebegn2003bs:\*  
10: genebegn2003cs:\*  
11: genebegn2003ds:\*  
12: genebegn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2679	100.0	2679	5 AAF30352	Aaf30352 Human chr
2	2133	79.6	2639	4 AAH14542	Aah14542 Human cDN
3	1995	74.5	1995	5 AAF89709	Aaf89709 Nucleotid
4	1618	60.4	3181	6 AAF89646	Aaf89646 Novel hum
5	1465	54.7	3138	4 AAH14556	Aah14556 Human cDN
6	1024	38.2	2259	4 AAH88903	Aah88903 Human pol
7	944	35.2	2448	11 ADM03388	Adm03388 Human cDN
8	765	28.6	1311	8 AAS25843	Aas25843 Human cDN
9	765	28.6	1311	8 ABX73184	Abx73184 Human nov
10	633	23.6	2186	10 ADA52592	Ada52592 Human cod
11	478	17.8	816	4 AAH06828	Aah06828 Human cDN
12	468	17.5	824	4 AAH06846	Aah06846 Human cDN
13	455	17.0	693	4 AAS26304	Aas26304 Human cDN
14	455	17.0	693	8 ABX73645	Abx73645 Human nov
15	426	15.9	518	4 AAH1859	Aah1859 Human cDN
16	285	10.7	575	3 AAA44336	Aaa44336 Human sec
17	267	10.0	357	4 AAS26143	Aas26143 Human cDN
18	267	10.0	357	8 ABX73484	Abx73484 Human nov
19	238	8.9	354	4 AAS26563	Aas26563 Human cDN
20	238	8.9	354	4 AAS26144	Aas26144 Human cDN
21	238	8.9	354	8 ABX73904	Abx73904 Human nov

22	238	8.9	354	8	ABX73485	Abx73485 Human nov
23	217	8.1	449	5	ACH26450	Ach26450 Human adu
24	159	6.3	1386	5	ABA16633	Ab16633 Human ner
25	139	5.2	825	4	AAK58374	Aak58374 Human lim
26	78	2.9	617	6	ABN60913	Abn60913 Human can
27	77	2.9	476	9	ACH14968	Ach14968 Human adu
28	60	2.2	60	6	ABN33472	Abn33472 Human spl
29	26	1.0	860	2	AAQ22007	Aaq22007 Sequence
30	25	0.9	1000	10	ADP45549	Adp45549 Arabidops
31	25	0.9	1065	3	AAAG6229	Aaag6229 cDNA enco
32	25	0.9	1963	6	AAQ39143	Aaq39143 Human tra
33	25	0.9	1963	10	ADD37442	Add37442 Human tra
34	25	0.9	2000	6	AAZ15532	Aaz15532 Arabidops
35	25	0.9	2000	8	ADA69051	Ada69051 Arabidops
36	25	0.9	2424	12	ADK67852	Adk67852 Phosphod
37	25	0.9	95769	8	ADK68659	Adk68659 Arabidops
38	24	0.9	199	6	ABL86644	Ab186644 Human ova
39	24	0.9	222	5	AAZ29141	Aaz29141 cDNA enco
40	24	0.9	222	5	AAZ29069	Aaz29069 cDNA enco
41	24	0.9	222	6	ABZ68281	Abz68281 cDNA enco
42	24	0.9	222	6	ABZ68209	Abz68209 cDNA enco
43	24	0.9	222	10	ADC25203	Adc25203 Human cDN
44	24	0.9	222	10	ADC25275	Adc25275 Human cDN
45	24	0.9	291	5	ABV20036	Abv20036 Human pro
46	24	0.9	413	5	ABV49795	Abv49795 Human pro
47	24	0.9	474	5	ACH23460	Ach23460 Human adu
48	24	0.9	525	5	ABV60912	Abv60912 Human pro
49	24	0.9	532	8	ACC60485	Acc60485 Rice leaf
50	24	0.9	532	10	ACP56929	Act56929 Rice leaf
51	24	0.9	703	6	ABO57317	Abg57317 Human col
52	24	0.9	1486	12	ADN75971	Adn75971 Human sof
53	24	0.9	2003	12	ADP22565	Adp22565 Sea-squir
54	24	0.9	2428	10	AAQ13316	Aaq13316 Human DMS
55	24	0.9	3157	2	AAQ13316	Aaq13316 Dufly rec
56	24	0.9	11869	4	AAZ36792	Aaz36792 Human car
57	24	0.9	11869	10	ADZ47486	Adz47486 Human car
58	24	0.9	178870	10	ADL13512	Adl13512 Osteocarth
59	24	0.9	349881	10	ADC86642	Adc86642 Human GPC
60	24	0.9	58	12	ADH05168	Adh05168 Gene poly
61	23	0.9	222	6	ABL87000	Ab187000 Human ova
62	23	0.9	242	4	AAI02577	Aai02577 Human rep
63	23	0.9	242	4	ABL97307	Ab197307 Human tes
64	23	0.9	374	6	ABK90577	Abk90577 Human tar
65	23	0.9	379	4	AAI87287	Aai87287 Human pol
66	23	0.9	410	4	AAI82951	Aai82951 Human pol
67	23	0.9	422	8	ABX38443	Abx38443 Bovine ES
68	23	0.9	469	9	ACH25805	Ach25805 Human adu
69	23	0.9	470	6	ABL93436	Ab193436 Arabidops
70	23	0.9	493	9	ACH32353	Ach32353 Human end
71	23	0.9	558	10	ACC55516	Acc55516 Rice endo
72	23	0.9	571	6	ABX09781	Abx09781 M. Incogn
73	23	0.9	592	8	ACC55401	Acc55401 Rice endo
74	23	0.9	601	12	ADQ63420	Adq63420 Transcript
75	23	0.9	636	2	AAQ38312	Aaq38312 ANS 0-13
76	23	0.9	726	4	AAI97185	Aai97185 Human neu
77	23	0.9	752	12	ADN88793	Adn88793 Nicotiana
78	23	0.9	759	6	ABQ55407	Abq55407 Human ova
79	23	0.9	799	6	AAI97521	Aai97521 Human neu
80	23	0.9	842	4	AAI94769	Aai94769 Human neu
81	23	0.9	922	10	ADC6496	Adc6496 Human GPC
82	23	0.9	942	4	AAK71033	Aak71033 Human fam
83	23	0.9	963	4	ABK43524	Abk43524 DNA enco
84	23	0.9	963	12	AD153911	Ad153911 cDNA enco
85	23	0.9	1125	6	ABN98318	Abn98318 Arabidops
86	23	0.9	1185	12	ADM94361	Adm94361 Soybean P
87	23	0.9	1222	2	AAZ83994	Aaz83994 Human BPC
88	23	0.9	1266	3	ABN81046	Abn81046 Shrimp po
89	23	0.9	1263	3	AAZ50385	Aaz50385 Mouse tae
90	23	0.9	1278	6	ABK61454	Abk61454 Human cDN
91	23	0.9	1279	10	AAI60064	Aai60064 Human PC4
92	23	0.9	1347	3	AAZ63330	Aaz63330 cDNA enco
93	23	0.9	1537	5	ABV25081	Abv25081 Human pro
94	23	0.9				







387	22	0.8	1114	9	ACG94813	Human sec	460	22	0.8	1114	9	ACD31673	Human sec
388	22	0.8	1114	9	ACD22532	Human sec	461	22	0.8	1114	9	ACD32594	Human sec
389	22	0.8	1114	9	ACFI5232	Human sec	462	22	0.8	1114	9	ACFI7560	Human sec
390	22	0.8	1114	9	ACCF93227	Human sec	463	22	0.8	1114	9	ACF07393	Human sec
391	22	0.8	1114	9	ACG92357	Human sec	464	22	0.8	1114	9	ACF20551	Human sec
392	22	0.8	1114	9	ACCF14004	Human sec	465	22	0.8	1114	9	ACF20858	Human sec
393	22	0.8	1114	9	ACFI4311	Human sec	466	22	0.8	1114	9	ACF21165	Human sec
394	22	0.8	1114	9	ACF09542	Human sec	467	22	0.8	1114	9	ACD47675	Human sec
395	22	0.8	1114	9	ACD68317	Novel hum	468	22	0.8	1114	9	ACF47657	Human sec
396	22	0.8	1114	9	ACD45833	Human sec	469	22	0.8	1114	9	ACF53490	Human sec
397	22	0.8	1114	9	ACD47982	Human sec	470	22	0.8	1114	9	ACD86825	Human sec
398	22	0.8	1114	9	ACD67713	CDNA enco	471	22	0.8	1114	9	ACH05073	CDNA enco
399	22	0.8	1114	9	ACF25521	Human sec	472	22	0.8	1114	9	ACF44570	Human sec
400	22	0.8	1114	9	ACF29205	Human sec	473	22	0.8	1114	9	ADAB1542	Human sec
401	22	0.8	1114	9	ACD84983	Human sec	474	22	0.8	1114	9	ACD22225	Human sec
402	22	0.8	1114	9	ACD84062	Human PRO	475	22	0.8	1114	9	ACD24572	Human sec
403	22	0.8	1114	9	ACD88053	Human sec	476	22	0.8	1114	9	ACD39775	Human sec
404	22	0.8	1114	9	ACCF30740	Human sec	477	22	0.8	1114	9	ACD40082	CDNA enco
405	22	0.8	1114	9	ACF32343	Human sec	478	22	0.8	1114	9	ACFI3390	Human sec
406	22	0.8	1114	9	ACH12003	CDNA enco	479	22	0.8	1114	9	ACF03192	Human sec
407	22	0.8	1114	9	ACH12310	CDNA enco	480	22	0.8	1114	9	ACF78663	Human sec
408	22	0.8	1114	9	ACD40702	Human sec	481	22	0.8	1114	9	ACFI1384	Human sec
409	22	0.8	1114	9	ACFI8174	Human sec	482	22	0.8	1114	9	ACF50727	Human sec
410	22	0.8	1114	9	ACF08621	Human sec	483	22	0.8	1114	9	ACF34222	Human sec
411	22	0.8	1114	9	ACF31422	Human sec	484	22	0.8	1114	9	ACD46447	Human sec
412	22	0.8	1114	9	ACF52262	Human sec	485	22	0.8	1114	9	ACD48289	Human sec
413	22	0.8	1114	9	ACD50131	Human sec	486	22	0.8	1114	9	ACF27670	Human sec
414	22	0.8	1114	9	ACF38834	Human sec	487	22	0.8	1114	9	ACF24542	Human sec
415	22	0.8	1114	9	ACF26749	Human sec	488	22	0.8	1114	9	ACD85597	Human sec
416	22	0.8	1114	9	ACF24849	Human sec	489	22	0.8	1114	9	ACD90202	Human sec
417	22	0.8	1114	9	ACF46429	Human sec	490	22	0.8	1114	9	ACD83755	Human PRO
418	22	0.8	1114	9	ACF27977	Human sec	491	22	0.8	1114	9	ACF49192	Human sec
419	22	0.8	1114	9	ACD89281	Human sec	492	22	0.8	1114	9	ACH07277	Human sec
420	22	0.8	1114	9	ACF63853	Human sec	493	22	0.8	1114	9	ACH07584	Human sec
421	22	0.8	1114	9	ACF60493	Human sec	494	22	0.8	1114	9	ACH08198	Human sec
422	22	0.8	1114	9	ACH12617	CDNA enco	495	22	0.8	1114	9	ACH11389	CDNA enco
423	22	0.8	1114	9	ACH10040	Human sec	496	22	0.8	1114	9	ACH11696	CDNA enco
424	22	0.8	1114	9	ACD03895	Human sec	497	22	0.8	1114	9	ACH10347	Human sec
425	22	0.8	1114	9	ACD10433	Human sec	498	22	0.8	1114	9	ACF01350	Human sec
426	22	0.8	1114	9	ACD12075	Human sec	499	22	0.8	1114	9	ACF40925	Human sec
427	22	0.8	1114	9	ACF42460	Human sec	500	22	0.8	1114	9	ACD24265	Human sec
428	22	0.8	1114	9	ACFI8481	Human sec	501	22	0.8	1114	9	ACD31366	Human sec
429	22	0.8	1114	9	ACF02271	Human sec	502	22	0.8	1114	9	ACFI1787	Human sec
430	22	0.8	1114	9	ACF21779	Human sec	503	22	0.8	1114	9	ACF32650	Human sec
431	22	0.8	1114	9	ACFI0463	Human sec	504	22	0.8	1114	9	ACF40311	Human sec
432	22	0.8	1114	9	ACF33915	Human sec	505	22	0.8	1114	9	ACF48271	Human sec
433	22	0.8	1114	9	ACF44877	Human sec	506	22	0.8	1114	9	ACF38220	Human sec
434	22	0.8	1114	9	ACD90509	Human sec	507	22	0.8	1114	9	ACF25156	Human sec
435	22	0.8	1114	9	ACD91132	Human sec	508	22	0.8	1114	9	ACF27056	Human sec
436	22	0.8	1114	9	ACF30433	Human sec	509	22	0.8	1114	9	ACF29512	Human sec
437	22	0.8	1114	9	ACD87132	Human sec	510	22	0.8	1114	9	ACD87746	Human sec
438	22	0.8	1114	9	ACF60186	Human sec	511	22	0.8	1114	9	ACF76207	Human sec
439	22	0.8	1114	9	ACF46736	Human sec	512	22	0.8	1114	9	ACF49499	Human sec
440	22	0.8	1114	9	ACF75593	Human sec	513	22	0.8	1114	9	ACF43956	Human sec
441	22	0.8	1114	9	ADAF9815	Human sec	514	22	0.8	1114	9	ACH06301	CDNA enco
442	22	0.8	1114	9	ACFI7253	Human sec	515	22	0.8	1114	9	ACH06608	CDNA enco
443	22	0.8	1114	9	ACF33007	Human sec	516	22	0.8	1114	9	ADAB3340	Human sec
444	22	0.8	1114	9	ACF08007	Human sec	517	22	0.8	1114	9	ACC92664	Human sec
445	22	0.8	1114	9	ACF08314	Human sec	518	22	0.8	1114	9	ACC93278	Human sec
446	22	0.8	1114	9	ACF40618	Human sec	519	22	0.8	1114	9	ACFI19323	Human sec
447	22	0.8	1114	9	ACF53797	Human sec	520	22	0.8	1114	9	ACD13014	Human sec
448	22	0.8	1114	9	ACD47061	Human sec	521	22	0.8	1114	9	ACF06472	Human sec
449	22	0.8	1114	9	ACF47964	Human sec	522	22	0.8	1114	9	ACC94506	Human sec
450	22	0.8	1114	9	ACF47350	Human sec	523	22	0.8	1114	9	ACC97934	Human sec
451	22	0.8	1114	9	ACF46122	Human sec	524	22	0.8	1114	9	ACC94199	Human sec
452	22	0.8	1114	9	ACD86211	Human sec	525	22	0.8	1114	9	ACF42153	Human sec
453	22	0.8	1114	9	ACF52569	Human sec	526	22	0.8	1114	9	ACD31059	Human sec
454	22	0.8	1114	9	ACF52876	Human sec	527	22	0.8	1114	9	ACD43088	CDNA enco
455	22	0.8	1114	9	ACF64869	Human sec	528	22	0.8	1114	9	ACD43395	CDNA enco
456	22	0.8	1114	9	ACF76514	Human sec	529	22	0.8	1114	9	ACFI14925	Human sec
457	22	0.8	1114	9	ACF61414	Human sec	530	22	0.8	1114	9	ACF01657	Human sec
458	22	0.8	1114	9	ACF61721	Human sec	531	22	0.8	1114	9	ACF31729	Human sec
459	22	0.8	1114	9	ACD30752	Human sec	532	22	0.8	1114	9	ACD67406	CDNA enco







971	22	0.8	2933	12	ADG91087
972	22	0.8	2933	12	ADG95228
973	22	0.8	2933	12	ADG93338
974	22	0.8	2933	12	ADG934919
975	22	0.8	2933	12	ADG92234
976	22	0.8	2933	12	ADG90535
977	22	0.8	2933	12	ADG91682
978	22	0.8	2933	12	ADG02261
979	22	0.8	2933	12	ADG22047
980	22	0.8	2933	12	ADG20117
981	22	0.8	2933	12	ADG98023
982	22	0.8	2933	12	ADG24240
983	22	0.8	2933	12	ADG98594
984	22	0.8	2933	12	ADG03425
985	22	0.8	2933	12	ADG99146
986	22	0.8	2933	12	ADG16731
987	22	0.8	2933	12	ADG05190
988	22	0.8	2933	12	ADG19457
989	22	0.8	2933	12	ADG13294
990	22	0.8	2933	12	ADG08351
991	22	0.8	2933	12	ADG15521
992	22	0.8	2933	12	ADG6919
993	22	0.8	2933	12	ADG06104
994	22	0.8	2933	12	ADG23688
995	22	0.8	2933	12	ADG03977
996	22	0.8	2933	12	ADG24878
997	22	0.8	2933	12	ADG07175
998	22	0.8	2933	12	ADG07727
999	22	0.8	2933	12	ADG55222
1000	22	0.8	2933	12	ADG60886

## ALIGNMENTS

## RESULT 1

AAAF30352 standard; cDNA, 2679 BP.

AC	AAAF30352;				
XX					
DT	14-MAY-2001 (first entry)				
XX					
DE	Human chr cDNA encoding checkpoint with FHA and ring finger protein.				
XX					
KW	Checkpoint with forkhead associated domain and ring finger; Chfr; human;				
KW	mitosis; cell cycle; tumour; diagnosis; antitumour; drug screening;				
KW	ubiquitin-protein ligase; ss.				
XX					
OS	Homo sapiens.				
XX					
FX	Key	Location/Qualifiers			
FT	CDS	91..2085			
FT		/tag= a			
FT	mutation	replace(1828,A)			
FT		/tag= b			
FT		/note= "alters codon GTG (Val) to ATG (Met) in chr from human cancer cell line U2OS"			
XX					
PN	WO200109150-A2.				
XX					
PD	08-FEB-2001.				
XX					
PF	14-JUN-2000; 2000WO-US016391.				
XX					
PR	29-JUL-1999; 99US-0146194P.				
XX					
PA	(WIST-) WISTAR INST ANATOMY & BIOLOGY.				
XX					
PI	Halazonectis T, Scolnick D;				
XX					
DR	WPI: 2001-182927/18.				
DR	P-PSDB; AAB20219.				

XX Novel nucleic acid sequence of mitotic checkpoint gene encoding a  
PT Checkpoint with forkhead-associated domain and ring finger protein, for  
PT diagnosing tumorigenic cells and in screening for anticancer drugs.  
XX  
PS Claim 2(a); Fig 4A-D; 85pp; English.

CC The present sequence is that of cDNA encoding the human mitotic  
CC checkpoint protein Chfr (see AAB20219) having a forkhead associated  
CC domain (FHA) and a ring finger domain. The protein is required for  
CC regulation of the transition of cells from prophase to metaphase during  
CC mitosis. It has ubiquitin-protein ligase activity. The Chfr checkpoint  
CC was evident in primary human cells, but was inactivated in 4 of 8 human  
CC cancer cell lines. In U2OS cells, a mutation was identified, which caused  
CC a Val to Met amino acid substitution in the highly conserved C-terminal  
CC Cys-rich region of the Chfr protein. In the absence of the Chfr  
CC checkpoint, cells subjected to mitotic stress condensed their chromosomes  
CC despite failing to separate their chromosomes. Chfr may monitor  
CC centrosome separation. Inactivation of the chfr gene in human cancer is  
CC theorized to underlie the increased sensitivity of cancer cells to  
CC antimetabolic drugs. Nucleic acids comprising the present sequence, or  
CC sequences encoding at least amino acids 31-103, 303-346, 476-641 (or  
CC their antisense sequences) are claimed. The chfr cDNA was isolated from  
CC an expressed sequence tag database sequence for cDNAs with FHA motifs.  
CC Claimed methods of determining the tumorigenic potential of a cell  
CC comprise examining the cell for the presence of chfr nucleic acid  
CC (absence or mutation indicating predisposition to tumorigenesis upon  
CC exposure to mitotic stress). Also claimed is a diagnostic agent, e.g. an  
CC antisense fragment of the present sequence, that binds to the chfr  
CC nucleic acid, and a diagnostic kit for detecting tumorigenic potential  
CC of a cell. A composition which inhibits the biological activity of Chfr  
CC may comprise a ligand selected from an antibody or its fragment. The Chfr  
CC inhibitor is used in a claimed method of retarding the growth of a cancer  
CC cells. Chfr polypeptides are useful in screening for drugs which can  
CC inhibit the activity of Chfr in a cancer cell, rendering the cell more  
CC sensitive to additional antitumour therapies

SQ Sequence 2679 BP, 650 A, 715 C, 764 G, 550 T; 0 U; 0 Other;

Query Match 100.0%; Score 2679; DB 5; Length 2679;  
Beet Local Similarity 100.0%; Pred. No. 0;  
Matches 2679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AAGAAATGGGACCGAGCGCGCAATGCTCTTGACAGCGCGCGCGCGCGGTTCCGG	60
DB	1	AAGAAATGGGACCGAGCGCGCAATGCTCTTGACAGCGCGCGCGCGCGGTTCCGG	60
QY	61	GTTCGGCGCGGGGCGGGATGGAATCCCGAATGAGCGCGCGCGAGAGGACGATCG	120
DB	61	GTTCGGCGCGGGGCGGGATGGAATCCCGAATGAGCGCGCGCGAGAGGACGATCG	120
QY	121	CCGCCCGCGAGCCCTGGGAGCGGCTCTGCTGGCGCGGAGAGGCGAGCGCGAC	180
DB	121	CCGCCCGCGAGCCCTGGGAGCGGCTCTGCTGGCGCGGAGAGGCGAGCGCGAC	180
QY	181	GTCTCTCTGAGAGAGCGGAGTGAACCTCGGCGCGGAGAGAGAGGTTGGACCTTC	240
DB	181	GTCTCTCTGAGAGAGCGGAGTGAACCTCGGCGCGGAGAGAGAGGTTGGACCTTC	240
QY	241	CCAGCAATAAAGTGTCTCTGAGATCACTGTGAATTTAGTGAAGAAATCAGAT	300
DB	241	CCAGCAATAAAGTGTCTCTGAGATCACTGTGAATTTAGTGAAGAAATCAGAT	300
QY	301	CAGGTGACATGGAAGATACAGCAACAGTGAACAGTGAATCAAGCTGAT	360
DB	301	CAGGTGACATGGAAGATACAGCAACAGTGAACAGTGAATCAAGCTGAT	360
QY	361	AAGAAGCAGATGCTCTTACAGATGCGGATGTCATCTCTGTTGTAAGAGAT	420
DB	361	AAGAAGCAGATGCTCTTACAGATGCGGATGTCATCTCTGTTGTAAGAGAT	420
QY	421	GAACCGAACAAGGTGACATCCTATGTAATTTAAGTAAGAAACAGGATGACA	480



Db 421 GAACGGAGCAACAGTGGATACCTCTATGAATCTTTAAGTGAAGCAAGGATGACA 480  
Qy 481 CAGAAATCCTTTAAGTAAACAAGAAATGTGTCCATGGAGCAAAAGTACTCGAGT 540  
Db 481 CAAAGAAATCTTTAAGTAAACAAGAAATGTGTTCATGGAGCAAAAGTACTCGAGT 540  
Qy 541 GCAGGTGACAGGGGAGGGGCGATCCCGGGTCCCTCGTGTGAGCCGCGACCTCAGGTG 600  
Db 541 GCAGGTGACAGGGGAGGGGCGATCCCGGGTCCCTCGTGTGAGCCGCGACCTCAGGTG 600  
Qy 601 TGTCTTGAAGAACCAAGCCATCAACATGCAAGCTCAGACCTCTTCCACAGCCTCGGC 660  
Db 601 TGTCTTGAAGAACCAAGCCATCAACATGCAAGCTCAGACCTCTTCCACAGCCTCGGC 660  
Qy 661 TCTTCAACGAGGCTTCTCTGCAAGGGGAGAGGCTTCTTCAAGTGTGGGTCTGGGGGT 720  
Db 661 TCTTCAACGAGGCTTCTCTGCAAGGGGAGAGGCTTCTTCAAGTGTGGGTCTGGGGGT 720  
Qy 721 GGTGGATCTCCCTTAAAGAAAGTGTCCCTGTGGCAAGTGAAGTCTCAGGCTT 780  
Db 721 GGTGGATCTCCCTTAAAGAAAGTGTCCCTGTGGCAAGTGAAGTCTCAGGCTT 780  
Qy 781 GCTCAGCTCTCCCAAGCAAGAAAGCTGCGTCTTCTGTGGAACCCAGAGTCAAG 840  
Db 781 GCTCAGCTCTCCCAAGCAAGAAAGCTGCGTCTTCTGTGGAACCCAGAGTCAAG 840  
Qy 841 GAGGATTTGAGAGCCGTGAAAGAAATGAGAGAGATGGGACCTTGAACCTGAGCGG 900  
Db 841 GAGGATTTGAGAGCCGTGAAAGAAATGAGAGAGATGGGACCTTGAACCTGAGCGG 900  
Qy 901 CAGTTTGGTCCGACAAACCGGTGAAGAAATGCCAAACCGTCCACAGAGAGCTCAGAGCA 960  
Db 901 CAGTTTGGTCCGACAAACCGGTGAAGAAATGCCAAACCGTCCACAGAGAGCTCAGAGCA 960  
Qy 961 GCGGCTGGAGAGCCAGCAAGATGAGAGAGAGCGCTGACATGATCAATCTGCAAGAGCTG 1020  
Db 961 GCGGCTGGAGAGCCAGCAAGATGAGAGAGAGCGCTGACATGATCAATCTGCAAGAGCTG 1020  
Qy 1021 CTGCAAGATCTGCGTGAATTTGCAAGCCCTGATGACACAGTTCGCGCGCTTCTACTCG 1080  
Db 1021 CTGCAAGATCTGCGTGAATTTGCAAGCCCTGATGACACAGTTCGCGCGCTTCTACTCG 1080  
Qy 1081 GGTGTGATGAGAGCTCTGCTCTGTGTCTTACCTGACCTGCGCGCTGAGCGAGTCTGT 1140  
Db 1081 GGTGTGATGAGAGCTCTGCTCTGTGTCTTACCTGACCTGCGCGCTGAGCGAGTCTGT 1140  
Qy 1141 AAAAAACACATCTTCAACAACTCTGTGAAGACATACCTCATCCAGATCCAGCAAGAGT 1200  
Db 1141 AAAAAACACATCTTCAACAACTCTGTGAAGACATACCTCATCCAGATCCAGCAAGAGT 1200  
Qy 1201 CGCAGTGAAGAAATGTGCAAAATGATGATGAGCCAGAAATTAATCACTCAAGACATGCTG 1260  
Db 1201 CGCAGTGAAGAAATGTGCAAAATGATGATGAGCCAGAAATTAATCACTCAAGACATGCTG 1260  
Qy 1261 CAGCCCAAAGTCAAGGCGGTCTTTTCTGATGAAGAGGAGTCAAGAGAGCTGCTGAG 1320  
Db 1261 CAGCCCAAAGTCAAGGCGGTCTTTTCTGATGAAGAGGAGTCAAGAGAGCTGCTGAG 1320  
Qy 1321 CTGTCAAGCTTGAAGATGATCTCAGACATTTAGCCAGCATACGTCTGTGCTCGGAG 1380  
Db 1321 CTGTCAAGCTTGAAGATGATCTCAGACATTTAGCCAGCATACGTCTGTGCTCGGAG 1380  
Qy 1381 TGTCTCAAGTGAAGAGGAGGAGGAGGAGCTTCCCACTGCGCAGACCCAGAGGCGAG 1440  
Db 1381 TGTCTCAAGTGAAGAGGAGGAGGAGGAGCTTCCCACTGCGCAGACCCAGAGGCGAG 1440  
Qy 1441 CCAAGAGCCCAACAGGCGCTGAGGAGTGAACCTTCAAGCTCGGTCAAGCTGAGCAAGCA 1500  
Db 1441 CCAAGAGCCCAACAGGCGCTGAGGAGTGAACCTTCAAGCTCGGTCAAGCTGAGCAAGCA 1500  
Qy 1501 GTTCAAGATTAAGTGTGCTCTGCAAGAGAGCAAGCCCTGTGACCTGTGCTTCCAG 1560  
Db 1501 GTTCAAGATTAAGTGTGCTCTGCAAGAGAGCAAGCCCTGTGACCTGTGCTTCCAG 1560

Qy 1561 CCAATGCCCAACGAGAGGAGGAGCCGAGACAGAGCCCGGTGTGCTGCTTCAAGAGT 1620  
Db 1561 CCAATGCCCAACGAGAGGAGGAGCCGAGACAGAGCCCGGTGTGCTGCTTCAAGAGT 1620  
Qy 1621 GCGGTCTGCTGAGAGCTTCTCTGCAAGCTTCACTGAGGAGCTGACCCGAGACCGGTGTAC 1680  
Db 1621 GCGGTCTGCTGAGAGCTTCTCTGCAAGCTTCACTGAGGAGCTGACCCGAGACCGGTGTAC 1680  
Qy 1681 GGTGCTGAGCCCGCTTGTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1740  
Db 1681 GGTGCTGAGCCCGCTTGTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1740  
Qy 1741 AACAACAAGCTACAGAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1800  
Db 1741 AACAACAAGCTACAGAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 1800  
Qy 1801 TGAAGAAACATGTTGAACCGAGAGCTCTGAGCTCTCAGAGGAGGAGTGTCTGCTGTCT 1860  
Db 1801 TGAAGAAACATGTTGAACCGAGAGCTCTGAGCTCTCAGAGGAGGAGTGTCTGCTGTCT 1860  
Qy 1861 GATTACAGAGTCAACGAGAGCACCGTCTGTGTTACTGTGTGCTGCTGCTGCTGCTGCT 1920  
Db 1861 GATTACAGAGTCAACGAGAGCACCGTCTGTGTTACTGTGTGCTGCTGCTGCTGCTGCTGCT 1920  
Qy 1921 GAGCTGACCTTACAGATGAGGAGAGACATCTCTGCTTCCGAGTTGCTGCTGCTGCTGCT 1980  
Db 1921 GAGCTGACCTTACAGATGAGGAGAGACATCTCTGCTTCCGAGTTGCTGCTGCTGCTGCT 1980  
Qy 1981 TCCGCTCTGAGCTGCTACTGAGGAGCGGTAACTGCGGACTGAGGAGGAGGAGGAGGAG 2040  
Db 1981 TCCGCTCTGAGCTGCTACTGAGGAGCGGTAACTGCGGACTGAGGAGGAGGAGGAGGAG 2040  
Qy 2041 ATGAATTCATATCTGTGAACAGACAGAGGTTCAAAATCTAAGATCCAGAGGCTT 2100  
Db 2041 ATGAATTCATATCTGTGAACAGACAGAGGTTCAAAATCTAAGATCCAGAGGCTT 2100  
Qy 2101 GAGCAGCTTTCAGACCTGAGAGGAGAGAGAGCTGTTTAAATACAGAGACAGAGAG 2160  
Db 2101 GAGCAGCTTTCAGACCTGAGAGGAGAGAGAGCTGTTTAAATACAGAGACAGAGAG 2160  
Qy 2161 TCAAGGTGTTTCAACAGCCCTGAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220  
Db 2161 TCAAGGTGTTTCAACAGCCCTGAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220  
Qy 2221 GGTGACTCTTCTGTGAGAGCTTCTTAACTGAGAGAGCTTCCCAAGAGCTTCCGAGG 2280  
Db 2221 GGTGACTCTTCTGTGAGAGCTTCTTAACTGAGAGAGCTTCCCAAGAGCTTCCGAGG 2280  
Qy 2281 CCGCAGCCCGCTCTCTGTGAGAGCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340  
Db 2281 CCGCAGCCCGCTCTCTGTGAGAGCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340  
Qy 2341 CGAAGCCTTCTGTGAACATGCGGCGCTGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 2400  
Db 2341 CGAAGCCTTCTGTGAACATGCGGCGCTGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 2400  
Qy 2401 ATTTCCGAACATCAAGTTAAAGCAGAGAGCTGTTTCAAGAAAGTTTCAAGGAGAGAG 2460  
Db 2401 ATTTCCGAACATCAAGTTAAAGCAGAGAGCTGTTTCAAGAAAGTTTCAAGGAGAGAG 2460  
Qy 2461 GGTGAGTTTATCAAAACATGTTTCAAGAGAGAGAGAGAGAGTTTCAAGCTTACAGAG 2520  
Db 2461 GGTGAGTTTATCAAAACATGTTTCAAGAGAGAGAGAGAGAGTTTCAAGCTTACAGAG 2520  
Qy 2521 CGTACACATATCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Db 2521 CGTACACATATCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Qy 2581 GGTGAGTTTATCTTCTTAAATGAGTTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2581 GGTGAGTTTATCTTCTTAAATGAGTTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640

QY 2641 ATTATATATTACAAATTTGACAAAAA 2679  
DB 2641 ATTATATATTACAAATTTGACAAAAA 2679

RESULT 2  
AAH14542  
ID AAH14542 standard; cDNA; 2639 BP.

XX AAH14542;  
XX  
XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:12099.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.

XX BP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INSR.

XX Oca T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR MPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
PT length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.

XX Claim 8; SEQ ID NO 12099; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primer sets are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

XX Sequence 2639 BP; 620 A; 706 C; 760 G; 553 T; 0 U; 0 Other;

Query Match 79.6%; Score 2133; DB 4; Length 2639;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 529 GATACCTCAGGTGACAGTGCAGAGGCGAGGCGATCCCGGGTCTCTCCGTGCGCC 588  
DB 481 GATACCTCAGGTGACAGTGCAGAGGCGAGGCGATCCCGGGTCTCTCCGTGCGCC 540  
QY 589 GCCACTCAGGTGTGCTTTGAGGAAACCAAGCCATTAACATGAGTCAAGCTCTTCC 648  
DB 541 GCCACTCAGGTGTGCTTTGAGGAAACCAAGCCATTAACATGAGTCAAGCTCTTCC 600  
QY 649 ACAGCCTGGGCTCTTCCAGGAGGCTCTCTGAGGCGAGAGCGTTCCTCAAGTTGT 708  
DB 601 ACAGCCTGGGCTCTTCCAGGAGGCTCTCTGAGGCGAGAGCGTTCCTCAAGTTGT 660  
QY 709 GGGTCTGGGGGTGTGGCATCTCCCTTAAGAAATGATGCTCTGTGGCAAGTATGA 768  
DB 661 GGGTCTGGGGGTGTGGCATCTCCCTTAAGAAATGATGCTCTGTGGCAAGTATGA 720  
QY 769 GTTCTCAGCTTTGCTCAGCTCTCTCCAGACAGAAAGATCTGGCTTTTCTGTTGAA 828  
DB 721 GTTCTCAGCTTTGCTCAGCTCTCTCCAGACAGAAAGATCTGGCTTTTCTGTTGAA 780  
QY 829 CCCAGATCAGAGAGATTTGAGAGCCGCTGAAGAAATGAGAGATGGGACCTT 888  
DB 761 CCCAGATCAGAGAGATTTGAGAGCCGCTGAAGAAATGAGAGATGGGACCTT 840  
QY 889 GACTTGAACGGGCACTGTTGTGTGACACACCGCTGAGAAATGCCCAACCGTCCAGAG 948  
DB 841 GACTTGAACGGGCACTGTTGTGTGACACACCGCTGAGAAATGCCCAACCGTCCAGAG 900  
QY 949 GACGTCAAGAGCAGCGCGTGGGAGAGCAGAAAGATGAGAGAGACCTGATGATCATC 1008  
DB 901 GACGTCAAGAGCAGCGCGTGGGAGAGCAGAAAGATGAGAGAGACCTGATGATCATC 960  
QY 1009 TGCAGGACCTGTGTCAGACATGCTGATGATTTGACAGCCCTGATGACACGTTCTGCG 1068  
DB 961 TGCAGGACCTGTGTCAGACATGCTGATGATTTGACAGCCCTGATGACACGTTCTGCG 1020  
QY 1069 GTTCTCTCTCTGGGCTGGATGAGGCGCTGTCTGTCTCTCACTGCGCTGTCCGCTG 1128  
DB 1021 GTTCTCTCTCTGGGCTGGATGAGGCGCTGTCTGTCTCTCACTGCGCTGTCCGCTG 1080  
QY 1129 GAGCGGATCTGTAAACCAACATCTCTCAACACCTCGGAGAGCATCTCATTCAGCAT 1188  
DB 1081 GAGCGGATCTGTAAACCAACATCTCTCAACACCTCGGAGAGCATCTCATTCAGCAT 1140  
QY 1189 CCAGCAAGAGTGCAGTGAAGAAAGATGTGCAAAATATGATGCCAGGATTAATCACT 1248  
DB 1141 CCAGCAAGAGTGCAGTGAAGAAAGATGTGCAAAATATGATGCCAGGATTAATCACT 1200  
QY 1249 CAAGCATGCTGCAGCCCAAGTCAAGGCGTCTTTTCTGATGAAAGAGAGTTCAGAG 1308  
DB 1201 CAAGCATGCTGCAGCCCAAGTCAAGGCGTCTTTTCTGATGAAAGAGAGTTCAGAG 1260  
QY 1309 GACCTGTGAGCTGTGCAAGTGAAGTGAAGTCTCTGACATTAAGCAGCATATACGTC 1368  
DB 1261 GACCTGTGAGCTGTGCAAGTGAAGTGAAGTCTCTGACATTAAGCAGCATATACGTC 1320  
QY 1369 GTGTGCGGCAGTGTCTGATGACAGAGCGGCGCAGCTCTCCCACTGCCAGCA 1428  
DB 1321 GTGTGCGGCAGTGTCTGATGACAGAGCGGCGCAGCTCTCCCACTGCCAGCA 1380  
QY 1429 CCGAGAGGCGAGGAGAGGAGCCCAAGGCGCTGGGGAGATGACCTTCACAGTCCGTACG 1488  
DB 1381 CCGAGAGGCGAGGAGAGGAGCCCAAGGCGCTGGGGAGATGACCTTCACAGTCCGTACG 1440  
QY 1489 CTGACGACAGAGTCCAGATTAAGTGTGCTCTGCAAGGAAACCAAGCCTGTGACAC 1548  
DB 1441 CTGACGACAGAGTCCAGATTAAGTGTGCTCTGCAAGGAAACCAAGCCTGTGACAC 1500  
QY 1549 TGTGCTTCCAGCCCATGCGGACCGAGAGCGGAGCGGAGCAGACCCGCTGTGCGCC 1608  
DB 1501 TGTGCTTCCAGCCCATGCGGACCGAGAGCGGAGCGGAGCAGACCCGCTGTGCGCC 1560



QY 391 GATGTCATCTACTTGGTGTACAGAAAGATGAACCGGAACACAGCTGGCATACCTCTAT 450  
DB 301 GAAGTCATCTACTTGGTGTACAGAAAGATGAACCGGAACACAGCTGGCATACCTCTAT 360  
QY 451 GAATCTTTAGTAAAGCAAGGCGATGACACAGAAATCCCTTTAGCTAAACAGAAAT 510  
DB 361 GAATCTTTAGTAAAGCAAGGCGATGACACAGAAATCCCTTTAGCTAAACAGAAAT 420  
QY 511 GTGTTCATGGGACCAAGATACCTCAGGTGCAAGGTGCAAGGCGAGGCGCATCCCGG 570  
DB 421 GTGTTCATGGGACCAAGATACCTCAGGTGCAAGGTGCAAGGCGAGGCGCATCCCGG 480  
QY 571 GTTCCTCCCTGTCGCGCGCCCACTCAGGTGTGCTTTGAGGAACCAAGCCATCAATG 630  
DB 481 GTTCCTCCCTGTCGCGCGCCCACTCAGGTGTGCTTTGAGGAACCAAGCCATCAATG 540  
QY 631 AGGTGACACTCTTCCCAAGGCTGAGGCTCTTCCAGGAGGCTCTCTGACGGGCGA 690  
DB 541 AGGTGACACTCTTCCCAAGGCTGAGGCTCTTCCAGGAGGCTCTCTGACGGGCGA 600  
QY 691 GAGCGTCTCTCAAGTTGTGGGTCTGGGGGTGTGGCATCTCCCTAAAGAAATGCTCC 750  
DB 601 GAGCGTCTCTCAAGTTGTGGGTCTGGGGGTGTGGCATCTCCCTAAAGAAATGCTCC 660  
QY 751 TCTGTGGCAAGTATGATGATCTCCAGCTTTGCTCAGCTCTCCCAACAGAAAGCTGCG 810  
DB 661 TCTGTGGCAAGTATGATGATCTCCAGCTTTGCTCAGCTCTCCCAACAGAAAGCTGCG 720  
QY 811 TCTTTTCTGCTGTGGAAACCCAGATCAGAGAGATTTTGAAGCCGTGAAGAAATG 870  
DB 721 TCTTTTCTGCTGTGGAAACCCAGATCAGAGAGATTTTGAAGCCGTGAAGAAATG 780  
QY 871 AAGAGAGATGGGACCTTGAACCTGAACGGGCAAGTTGTGTGCAACACCGGTGAAT 930  
DB 781 AAGAGAGATGGGACCTTGAACCTGAACGGGCAAGTTGTGTGCAACACCGGTGAAT 840  
QY 931 GCCCAAAACCTGTCACAGAGACCTCAGAGAGAGGCGGCTGGGAAGCCAGAAAGTGAAG 990  
DB 841 GCCCAAAACCTGTCACAGAGACCTCAGAGAGAGGCGGCTGGGAAGCCAGAAAGTGAAG 900  
QY 991 ACGGTGACATGATCATCTGCGAGACCTGCTGCAAGACTGCGTGAAGTTGAGCCCTG 1050  
DB 901 ACGGTGACATGATCATCTGCGAGACCTGCTGCAAGACTGCGTGAAGTTGAGCCCTG 960  
QY 1051 ATGCAACAGTCTGCGGCGCTTGTACTGCGGCGTGTGATGAGGAGCTGCTCCGTGCT 1110  
DB 961 ATGCAACAGTCTGCGGCGCTTGTACTGCGGCGTGTGATGAGGAGCTGCTCCGTGCT 1020  
QY 1111 AACTGCGCGTGTCCCTGAGAGGCGATCTGTAAANACCAATCCTCAACAACTCTGTGAA 1170  
DB 1021 AACTGCGCGTGTCCCTGAGAGGCGATCTGTAAANACCAATCCTCAACAACTCTGTGAA 1080  
QY 1171 GATAACCTCATCAGATCAGACAGAGATCGCAGTGAAGAGATGTGCAAGATGTGAT 1230  
DB 1081 GATAACCTCATCAGATCAGACAGAGATCGCAGTGAAGAGATGTGCAAGATGTGAT 1140  
QY 1231 GCCAGAAATTAATCACTCAAGCATGTGACAGCCCAAGATCAAGGCGGTCTTTTCTGAT 1290  
DB 1141 GCCAGAAATTAATCACTCAAGCATGTGACAGCCCAAGATCAAGGCGGTCTTTTCTGAT 1200  
QY 1291 GAAAGAGGAGTTCAGAGAGACCTGCTGAGCTGTCAAGTGTGACAGTGTCTCAAGC 1350  
DB 1201 GAAAGAGGAGTTCAGAGAGACCTGCTGAGCTGTCAAGTGTGACAGTGTCTCAAGC 1260  
QY 1351 ATTAGCAGCCATACGTGTGTGCGCGAGTGTCTGAGTACAGAAAGCAGGCGGCGAG 1410  
DB 1261 ATTAGCAGCCATACGTGTGTGCGCGAGTGTCTGAGTACAGAAAGCAGGCGGCGAG 1320  
QY 1411 CCTCCCACTGCGCAGACCCGAGGCGAGCCAGAGGCCCAAGGCTCTGGGGGATGCA 1470  
DB 1321 CCTCCCACTGCGCAGACCCGAGGCGAGCCAGAGGCCCAAGGCTCTGGGGGATGCA 1380

QY 1471 CCTCCCACTGCGCAGACCCGAGGCGAGCCAGAGGCCCAAGGCTCTGGGGGATGCA 1530  
DB 1381 CCTCCCACTGCGCAGACCCGAGGCGAGCCAGAGGCCCAAGGCTCTGGGGGATGCA 1440  
QY 1531 AGCCAGCCCTGTGACCTGCTGTCCAGGCCCATAGCCGAGCGGAGCGGAGCGGAG 1590  
DB 1441 AGCCAGCCCTGTGACCTGCTGTCCAGGCCCATAGCCGAGCGGAGCGGAGCGGAG 1500  
QY 1591 CAGAACCGCGGTGTGCGCCCTCAGAGAGTGTGGGTCTGCTGCAAGCTTTCTGCACTG 1650  
DB 1501 CAGAACCGCGGTGTGCGCCCTCAGAGAGTGTGGGTCTGCTGCAAGCTTTCTGCACTG 1560  
QY 1651 TACTGGGCTGACACCCGAGCCGAGCTGTCAAGGCTGTGCGCCGCTTTTGTGAGCTCAAC 1710  
DB 1561 TACTGGGCTGACACCCGAGCCGAGCTGTCAAGGCTGTGCGCCGCTTTTGTGAGCTCAAC 1620  
QY 1711 CTGGGTGACAAAGTGTCTGAGACGCGCTGTGAACAACAACAGCTACAGATCAGATCTCG 1770  
DB 1621 CTGGGTGACAAAGTGTCTGAGACGCGCTGTGTGAACAACAACAGCTACAGATCAGATCTCG 1680  
QY 1771 AAGAAATTAACCTGGCAACCAAGGTTTGAATGAGAAAAATGTTGACCGAGACCTCGTG 1830  
DB 1681 AAGAAATTAACCTGGCAACCAAGGTTTGAATGAGAAAAATGTTGACCGAGACCTCGTG 1740  
QY 1831 GCTTCCAGCCGAGAGTGTCTGTGTGTGATTAACAAGTCAAGGAGACACCGTTCTG 1890  
DB 1741 GCTTCCAGCCGAGAGTGTGTGTGTGATTAACAAGTCAAGGAGACACCGTTCTG 1800  
QY 1891 TGTATCTGCTGTGAGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGAGAACTT 1950  
DB 1801 TGTATCTGCTGTGAGCTGCGAGCTTCCGTGAGCTGACCTATCAGTATCGGAGAACTT 1860  
QY 1951 CCTGCTTCGAGTGGCCAGAGTGGCGGTAAATATCCGCTCTGACTGTGAGGCGCTTAAC 2010  
DB 1861 CCTGCTTCGAGTGGCCAGAGTGGCGGTAAATATCCGCTCTGACTGTGAGGCGCTTAAC 1920  
QY 2011 TGCCGCACTCAGGTAAAGCTCACACAGCCATGAAATCAATCATATCTGTGAACAGCA 2070  
DB 1921 TGCCGCACTCAGGTAAAGCTCACACAGCCATGAAATCAATCATATCTGTGAACAGCA 1980  
QY 2071 AGGTTCAAAAATAA 2085  
DB 1981 AGGTTCAAAAATAA 1995  
  
RESULT 4  
ABNS9646  
ID ABNS9646 standard; cDNA, 3181 BP.  
XX  
AC ABNS9646;  
XX  
DT 28-JUN-2002 (first entry)  
XX  
DE Novel human coding sequence SEQ ID NO: 57.  
XX  
KW Human; antianemic; vulnerary; antinflammatory; immunomodulator;  
XX antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; BST;  
XX expressed sequence tag; gene; ss.  
OS Homo sapiens.  
XX  
FN WO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 10-SEP-2001; 2001MO-US026015.  
XX  
PR 11-SEP-2000; 2000US-00659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX MPI, 2002-292408/33.  
DR P-PSDB; ABB97233.  
XX An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
PS Claim 1, SEQ ID NO 57, 509pp, English.  
XX  
XX The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a coding sequence of the  
CC invention  
XX  
SQ Sequence 3181 BP, 756 A; 830 C; 879 G; 716 T; 0 U; 0 Other;  
Query Match 60.4%; Score 1618; DB 6; Length 3181;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1948; Conservative 0; Mismatches 3; Indels 2; Gaps 1;  
QY 709 GGGTCTGGGGGTTGGTCATCTCCCTTAAGAAAGTGTCTCTGTGGCAATGATGA 768  
DB 591 GGGTCTGGGGGTTGGTCATCTCCCTTAAGAAAGTGTCTCTGTGGCAATGATGA 650  
QY 769 GTTCTCAAGCTTTGGCTTCAAGTCTTCCAGACAGAAAGCTGCTTTTGTGTGGA 828  
DB 651 GTTCTCAAGCTTTGGCTTCAAGTCTTCCAGACAGAAAGCTGCTTTTGTGTGGA 710  
QY 829 CCCCAGATCAGAGAGATTGGAGCCCGTGAAGAAATGAGAGATGGGAGCTT 888  
DB 711 CCCCAGATCAGAGAGATTGGAGCCCGTGAAGAAATGAGAGATGGGAGCTT 770  
QY 889 GACCTGAACGGGACGTTGTGTGCAACAACCGGTGAATGCCAACCCTCAAG 948  
DB 771 GACCTGAACGGGACGTTGTGTGCAACAACCGGTGAATGCCAACCCTCAAG 830  
QY 949 GACCTGAACGGGACGTTGTGGAAGCCAGACAGATGAGAGAGCTGATCATCATC 1008  
DB 831 GACCTGAACGGGACGTTGTGGAAGCCAGACAGATGAGAGAGCTGATCATCATC 890  
QY 1009 TGCAGAGACCTGCTGCAAGCTGTGAGTTTGCAGCCCTGCATGCAACAGTTCTGGCG 1068  
DB 891 TGCAGAGACCTGCTGCAAGCTGTGAGTTTGCAGCCCTGCATGCAACAGTTCTGGCG 950  
QY 1069 GCTTGTCTACTCGGAGCTGATGAGAGCGCTCGCTGTGTCTTACTGCTGCTGCTGCTG 1128  
DB 951 GCTTGTCTACTCGGAGCTGATGAGAGCGCTCGCTGTGTCTTACTGCTGCTGCTGCTG 1010  
QY 1129 GAGCGGATCTGTAAAAACCAATCTTCAACAACCTCTGGAAGATACCTCATCAGCAT 1188  
DB 1011 GAGCGGATCTGTAAAAACCAATCTTCAACAACCTCTGGAAGATACCTCATCAGCAT 1070  
QY 1189 CAGACAGAGTGTGAGTGAAGAAAGTGTGCAAGTGTGATGCGAGGAATTAATCACT 1248  
DB 1071 CAGACAGAGTGTGAGTGAAGAAAGTGTGCAAGTGTGATGCGAGGAATTAATCACT 1130  
QY 1249 CAGACAGTGTGAGTGAAGTCAAGCGGTCTTTTCTGATGAAGAAAGGAGTTCAAG 1308  
DB 1131 CAGACAGTGTGAGTGAAGTCAAGCGGTCTTTTCTGATGAAGAAAGGAGTTCAAG 1190  
QY 1309 GACCTGTCTGAGCTGTGAGAGCTTGAAGAGTCTTCAAGCATTTAGCCAGCATAGCTC 1368  
DB 1191 GACCTGTCTGAGCTGTGAGAGCTTGAAGAGTCTTCAAGCATTTAGCCAGCATAGCTC 1250  
QY 1369 GTGTGCGGAGAGTGTCTGAGTACAGAAAGCAGCGCGAGGCTCCCACTGCCAGCA 1428

DB 1251 GTGTGCGGAGAGTGTCTGAGTACAGAAAGCAGCGCGAGGCTCCCACTGCCAGCA 1310  
QY 1429 CCGAGAGGCGAGCAGAGAGCCCAACAGGCTCTGGGGAGTACACCTTCAAGTCCGTCAAGC 1488  
DB 1311 CCGAGAGGCGAGCAGAGAGCCCAACAGGCTCTGGGGAGTACACCTCCCACTGCTCAAGC 1370  
QY 1489 CTGACAGACAGTGTGAGTGAAGTGTGCTGTGCAAGGAGGAGCCCTGTGCAAC 1548  
DB 1371 CTGACAGACAGTGTGAGTGAAGTGTGCTGTGCAAGGAGGAGCCCTGTGCAAC 1430  
QY 1549 TGTGCTTCCAGAGCCCAATGAGCCCAACAGGAGCAGAGCAGAGAGACCCGCTGTGAGC 1608  
DB 1431 TGTGCTTCCAGAGCCCAATGAGCCCAACAGGAGCAGAGAGACCCGCTGTGAGC 1490  
QY 1609 CTTCAAGAGTGTGAGTGTGCTGTGCAAGCTTCTGCACTGTGATCTGAGGCTGACCCG 1668  
DB 1491 CTTCAAGAGTGTGAGTGTGCTGTGCAAGCTTCTGCACTGTGATCTGAGGCTGACCCG 1550  
QY 1669 ACCGCTGTGAGTGTGCTGTGAGTGTGCTGTGCAAGTGTGATCTGAGGCTGACCCG 1728  
DB 1551 ACCGCTGTGAGTGTGCTGTGAGTGTGCTGTGCAAGTGTGATCTGAGGCTGACCCG 1610  
QY 1729 GACGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1788  
DB 1611 GACGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1670  
QY 1789 AGAGTGTGATGATGAAAAATCATGTTGACCGAGAGCTGTGAGCTTCCAGCGGAGAGT 1848  
DB 1671 AGAGTGTGATGATGAAAAATCATGTTGACCGAGAGCTGTGAGCTTCCAGCGGAGAGT 1730  
QY 1849 TTTCTGCTGTGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1908  
DB 1731 TTTCTGCTGTGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1790  
QY 1909 CCGAGCTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1968  
DB 1791 CCGAGCTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1850  
QY 1969 GTGCGGTGAACATCCGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2028  
DB 1851 GTGCGGTGAACATCCGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1910  
QY 2029 GCTCAACAGCCATGAAATTCATATCTGTGGAACAGACAGTGTGAAATTCATAGCA 2088  
DB 1911 GCTCAACAGCCATGAAATTCATATCTGTGGAACAGACAGTGTGAAATTCATAGCA 1970  
QY 2089 TCCAGAGGCTCTGAGAGCTTTCAGACCTGAGAGTGAAGAGACGTTTAAATTAACA 2148  
DB 1971 TCCAGAGGCTCTGAGAGCTTTCAGACCTGAGAGTGAAGAGACGTTTAAATTAACA 2030  
QY 2149 GAGACAGCAAGTCAAGAGTGTTCACAGCCCTGAGAGAGAGAGAGAGGAGTCTCCGA 2208  
DB 2031 GAGACAGCAAGTCAAGAGTGTTCACAGCCCTGAGAGAGAGAGAGAGGAGTCTCCGA 2090  
QY 2209 CAGGTCTCTGGGAGTGTGATCTGTGAGAGCTTTTAACTCTGATGAGAGCCCTCCCA 2268  
DB 2091 CAGGTCTCTGGGAGTGTGATCTGTGAGAGC--TTTAACTCTGATGAGAGCCCTCCCA 2148  
QY 2269 GAGCCCGGAGGCGAGAGCCGCTCTGAGTGAAGAGTGTGAGAGGAGTGTGAGTGTGAGTGTG 2328  
DB 2149 GAGCCCGGAGGCGAGAGCCGCTCTGAGTGAAGAGTGTGAGAGGAGTGTGAGTGTGAGTGTG 2208  
QY 2329 CAGACAGAGAGAGAGAGCTTTCTGTAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2388  
DB 2209 CAGACAGAGAGAGAGAGCTTTCTGTAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2268  
QY 2389 GCTCTTTTGTACATTTTCCGAAATCAAGTGAAGAGAGAGTGTGTTTCAAGAAAGT 2448  
DB 2269 GCTCTTTTGTACATTTTCCGAAATCAAGTGAAGAGAGTGTGTTTCAAGAAAGT 2328  
QY 2449 TCAAGGAGAGAGGAGGAGTATCAAAAACATGTTTCAAGAGAGAGAGATAGTTTAA 2508







Db 1460 CCTCAGCAGTGTGCGGTCTGCTGACGCTTTCTGCCACCTGTATCTGCGGTGACCCCG 1519  
 QY 1669 ACCGGCTGCTAAGGCTGCGCCCGCTTTTGTAGCTCAACCTGGGTGA CAAGTGTCTG 1728  
 Db 1520 ACCGGCTGCTAAGGCTGCGCCCGCTTTTGTAGCTCAACCTGGGTGA CAAGTGTCTG 1579  
 QY 1729 GACGGCGTGTGAACAACAAGCTACGAGTACAGACATCTGTAAGATTAATCTGCGAAC 1788  
 Db 1580 GACGGCGTGTGAACAACAAGCTACGAGTACAGACATCTGTAAGATTAATCTGCGAAC 1539  
 QY 1789 AGAGGTTTGCATGTAAGAAAATGTTGACCGAGAGCTCTGTGCTCTCCAGCGGGAGTG 1848  
 Db 1640 AGAGGTTTGCATGTAAGAAAATGTTGACCGAGAGCTCTGTGCTCTCCAGCGGGAGTG 1699  
 QY 1849 TTTCTGCTGTCTGATTAACAGATCAAGGAGACACCGTTGTGTATCTGCTGTGCGT 1908  
 Db 1700 TTTCTGCTGTCTGATTAACAGATCAAGGAGACACCGTTGTGTATCTGCTGTGCGT 1759  
 QY 1909 CCGAGCTTCCGTAGCTGACCTATCAGTATCGGCAAGAACATTCCTGCTCCGAGTTGCCA 1968  
 Db 1760 CCGAGCTTCCGTAGCTGACCTATCAGTATCAGCAAGAACATTCCTGCTCCGAGTTGCCA 1819  
 QY 1969 GTGGCGGTAAATCCGCTCTGACCTGCTACTGAGGCGGTAACTGCGCACTCAGGTGAA 2028  
 Db 1820 GTGGCGGTAAATCCGCTCTGACCTGCTACTGAGGCGGTAACTGCGCACTCAGGTGAA 1879  
 QY 2029 GCTCACCAGCCCATGAAATTCATCATCTGTGAAACAGACAGTTTCAAAAATAAGCA 2088  
 Db 1880 GCTCACCAGCCCATGAAATTCATCATCTGTGAAACAGACAGTTTCAAAAATAAGCA 1939  
 QY 2089 TCCAGAGCCCTGAGAGCTTTTCAAGCTGAGGTGAAGAGAGCTGTTTAAATATCA 2148  
 Db 1940 TCCAGAGCCCTGAGAGCTTTTCAAGCTGAGGTGAAGAGAGCTGTTTAAATATCA 1999  
 QY 2149 GAGACAGACGCTCAAGGTGTTTCAAGCCCTCTGAGGAGAGAGACGAGGCTCTCCGA 2208  
 Db 2000 GAGGCAAGACGCTCAAGGTGTTTCAAGCCCTCTGAGGAGAGAGACGAGGCTCTCCGA 2059  
 QY 2209 CAGGTGCTCTGGGTGATCTTTCTGTGAGACTTTTAACCTCTGATGAGACCTTCCCA 2268  
 Db 2060 CAGGTGCTCTGGGTGATCTTTCTGTGAGACTTTTAACCTCTGATGAGACCTTCCCA 2117  
 QY 2269 GAGCCCGGAGGCGGAGCCGCTCTCTGTGAGGCTGAGGAGGCTCTGTGTGACAT 2328  
 Db 2118 GAGCCCGGAGGCGGAGCCGCTCTCTGTGAGGCTGAGGAGGCTCTGTGTGACAT 2177  
 QY 2329 CAGCAGCAGAGAGAGCCCTTTCTGTAACTATGCGCGCTCTCCCGCAGAGGCGCAGTTT 2388  
 Db 2178 CAGCAGCAGAGAGAGCCCTTTCTGTAACTATGCGCGCTCTCCCGCAGAGGCGCAGTTT 2237  
 QY 2389 GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGCAAGATCTGTTTTCAGGAAAAGTT 2448  
 Db 2238 GCTCTTTTGTACATTTTCCGAACTACAGTTAAAGCAAGATCTGTTTTCAGGAAAAGTT 2297  
 QY 2449 TCAAGGAGAGAGGAGGAGTTTATCAAAAACATGTTTTCAGGAGAGGAGCATTAAGTTTA 2508  
 Db 2298 TCAAGGAGAGAGGAGGAGTTTATCAAAAACATGTTTTCAGGAGAGGAGCATTAAGTTTA 2357  
 QY 2509 CAGCCTACAGAGAGTACAAATATCTGCTGCTGAGGAAAACAAGCATTTTATCTATTT 2568  
 Db 2358 CAGCCTACAGAGAGTACAAATATCTGCTGCTGAGGAAAACAAGCATTTTATCTATTT 2417  
 QY 2569 TTTATTTTAAATAGTTTGTGCTTATCTTAAATTAAGATTTAAATGTCAAACTGTAGC 2628  
 Db 2418 TTTATTTTAAATAGTTTGTGCTTATCTTAAATTAAGATTTAAATGTCAAACTGTAGC 2477  
 QY 2629 ACAAAATATATATTTATATATTTACAAATTTAC 2661  
 Db 2478 ACAAAATATATATTTATATATTTACAAATTTAC 2510

RESULT 6  
 AA188903/c

ID AA188903 standard; cDNA; 2259 BP.  
 AC AA188903; c.  
 DT 06-NOV-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 8963.  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation; ss.  
 OS Homo sapiens.  
 EN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 PF 26-FEB-2001; 2001WO-US004927.  
 PR 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Dmanac RT;  
 DR WPI; 2001-514838/56.  
 DR P-PSDB; AA008972.  
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukaemia, inflammation and immune disorders.  
 XX  
 PS Claim 1; SEQ ID NO 8963; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AA19941-AA19984) and  
 CC the encoded proteins (AA00010-AA013910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pot\_sequences  
 XX  
 SQ Sequence 2259 BP; 531 A; 612 C; 615 G; 501 T; 0 U; 0 Other;  
 Query Match 38.2%; Score 1024; DB 4; Length 2259;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1304; Conservative 0; Mismatches 2; Indels 2; Gaps 1;  
 QY 1354 AGCCAGCCATPAGTCTGTGCGGCGAGTCTCTGATTAAGAGAGCGCGGACGCT 1413  
 Db 1965 AGCCAGCCATPAGTCTGTGCGGCGAGTCTCTGATTAAGAGAGCGCGGACGCT 1906  
 QY 1414 CCCCACTGCCAGACCCCGAGGCGGAGAGAGGCCCAAGGCCCTGTGGGAGATGACCC 1473  
 Db 1905 CCCCACTGCCAGACCCCGAGGCGGAGAGAGGCCCAAGGCCCTGTGGGAGATGACCC 1846  
 QY 1474 TCCAGCTCCGTAGCTGACAGACAGAGTCCAGATTAAGTGTGCTCTGCAAGAGAC 1533  
 Db 1845 TCCAGCTCCGTAGCTGACAGACAGAGTCCAGATTAAGTGTGCTCTGCAAGAGAC 1786  
 QY 1534 CAGCCCTGTGCACTGCTGCTTCCAGGCCATATGCCAGCCGAGAGAGCGGAGCAG 1593  
 Db 1785 CAGCCCTGTGCACTGCTGCTTCCAGGCCATATGCCAGCCGAGAGAGCGGAGCAG 1726  
 QY 1594 GACCCGCGTGTGCGCCCTCAGAGAGTGTGGGTCTGCTGCAAGCTTTCTGCCACTGTAC 1653

```

Db      1725 GACCCGCGTGTGCCCCCTTACGACAGTGTGCGTCTGAGCCCTTCTGCAACTGTGAC 1666
Qy      1654 TGGGGGCTGACCCCGGACCGGCTGCTACGGCTGCTGCGCCCGTTTGTAGTCAACTG 1713
Db      1665 TGGGGGCTGACCCCGGACCGGCTGCTACGGCTGCTGCGCCCGTTTGTAGTCAACTG 1606
Qy      1714 GGTGACAAAGTGTGTGACCGGCTGCTGAACAACAAGCTACGAGTCAAGTCTGAGG 1773
Db      1605 GGTGACAAAGTGTGTGACCGGCTGCTGAACAACAAGCTACGAGTCAAGTCTGAGG 1546
Qy      1774 AATTACTGGAACAAGAGTTTGAATGAGGAAAAATGATGTCGAGAGCCGCGGCT 1833
Db      1545 AATTACTGGAACAAGAGTTTGAATGAGGAAAAATGATGTCGAGAGCCGCGGCT 1486
Qy      1834 CTCACGCGGAGAGTGTCTGCTGTGATTAACAAGTACAGGAGACAACGTTGTGTGT 1893
Db      1485 CTCACGCGGAGAGTGTCTGCTGTGATTAACAAGTACAGGAGACAACGTTGTGTGT 1426
Qy      1894 TACTGTGTGCGCTGCGAGCTTCCGTGAGTGAACCTATCAGTATCGAGAACATTCCT 1953
Db      1425 TACTGTGTGCGCTGCGAGCTTCCGTGAGTGAACCTATCAGTATCGAGAACATTCCT 1366
Qy      1954 GCTTCGAGTGTGCAAGTGTGCGGCTTACATCCGCTCTGACTGCTAGCGGCGTAACTGC 2013
Db      1365 GCTTCGAGTGTGCAAGTGTGCGGCTTACATCCGCTCTGACTGCTAGCGGCGGCTAACTGC 1306
Qy      2014 CGCACTCAGGTGAAGCTCAACAAGCATGAATTCATCATCTGTGAACAGACAGG 2073
Db      1305 CGCACTCAGGTGAAGCTCAACAAGCATGAATTCATCATCTGTGAACAGACAGG 1246
Qy      2074 TTCAAAACTAGAGCTCCAGAGGCGCTGAGAGCTTTCAGCACTGGAAGTGAAGAGAGG 2133
Db      1245 TTCAAAACTAGAGCTCCAGAGGCGCTGAGAGCTTTCAGCACTGGAAGTGAAGAGAGG 1186
Qy      2134 TCTTTTAAATAACAGAGACAAGCAAGTCAAGTGTGTTTCAAGGCCCCCTGAGGAGAGG 2193
Db      1185 TCTTTTAAATAACAGAGACAAGCAAGTCAAGTGTGTTTCAAGGCCCCCTGAGGAGAGG 1126
Qy      2194 AGCGAGGCTCTCCGACAGGCTCTCTGGGGTGAATCTTCTGTGAGAGTCTTTTAACTCTGA 2253
Db      1125 AGCGAGGCTCTCCGACAGGCTCTCTGGGGTGAATCTTCTGTGAGAGCTTTTAACTCTGA 1068
Qy      2254 GTGAGACCTCCCGACAGAGCCCGGGGGGCGGAGCCCGCTCTGTGAGAGCTGGGAG 2313
Db      1067 GTGAGACCTCCCGACAGAGCCCGGGGGGCGGAGCCCGCTCTGTGAGAGCTGGGAG 1008
Qy      2314 GGTCTGTGTGATCAGACAGAGACAAGACCTTCTGTAAACAATGCGGCGCTCCGCG 2373
Db      1007 GGTCTGTGTGATCAGACAGAGACAAGACCTTCTGTAAACAATGCGGCGCTCCGCG 948
Qy      2374 GAGAGGGGAGATTTGCTCTTTTGTACATTTTCCGAACACTACAGTTAAAGCAAGTCTG 2433
Db      947 GAGAGGGGAGATTTGCTCTTTTGTACATTTTCCGAACACTACAGTTAAAGCAAGTCTG 888
Qy      2434 TTTTTCAGAAAAGTTTCAAGGAGAGGCAAGTTTCAAAAACATTTGTTCAAGAGAA 2493
Db      887 TTTTTCAGAAAAGTTTCAAGGAGAGGCAAGTTTCAAAAACATTTGTTCAAGAGAA 828
Qy      2494 GGGAGCATTAAGTTTACAGCTTACAGGACGTACACAATATCTGCTGTGAGGAAACACA 2553
Db      827 GGGAGCATTAAGTTTACAGCTTACAGGACGTACACAATATCTGCTGTGAGGAAACACA 768
Qy      2554 GCATTTTATCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2613
Db      767 GCATTTTATCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 708
Qy      2614 GTCAAACTGTAGCAAAATATATATTTTATTTTATTTTATTTTATTTTATTTTATTT 2661
Db      707 GTCAAACTGTAGCAAAATATATATTTTATTTTATTTTATTTTATTTTATTTTATTT 660

```

RESULT 7

```

ADMO3388
ID      ADMO3388 standard; cDNA; 2448 BP.
XX
AC      ADMO3388;
XX
DT      20-MAY-2004 (first entry)
XX
DE      Human cDNA of the invention SEQ ID NO:2073.
XX
KW      ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS      Homo sapiens.
XX
PN      EP1347046-A1.
XX
PD      24-SEP-2003.
XX
PF      12-APR-2002; 2002EP-00008400.
XX
PR      22-MAR-2002; 2002JP-00137785.
XX
PA      (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI      Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI      Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI      Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
DR      WPI; 2003-723558/69.
DR      P-Psdb; ADM05831.
XX
PT      New polynucleotides and polypeptides are useful in gene therapy, for
PT      developing a diagnostic marker or medicines for regulating their
PT      expression and activity, or as a target of gene therapy.
XX
PS      Claim 1; SEQ ID NO 2073; 305bp; English.
XX
CC      The invention relates to a novel human polynucleotide and the encoded
CC      polypeptide. A polynucleotide of the invention may have a use in gene
CC      therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC      as a primer for synthesizing the polynucleotide or as a probe for
CC      detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC      useful in gene therapy, for developing a diagnostic marker or medicines
CC      for regulating their expression and activity, or as a target of gene
CC      therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC      are useful as pharmaceutical agents. The present sequence represents a
CC      cDNA sequence of the invention.
XX
SQ      Sequence 2448 BP; 618 A; 624 C; 661 G; 545 T; 0 U; 0 Other;
XX
Query Match      35.2%; Score 944; DB 11; Length 2448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 994; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      708 TGGGTCTGGGGGTGTGTCATCTCCCTTAAAGAGTGTCTCTGTGCAAGTATGA 767
Db      357 TGGGTCTGGGGGTGTGTCATCTCCCTTAAAGAGTGTCTCTGTGCAAGTATGA 416
Qy      768 AGTCTCAGCTTTGCTTCACTTCCAGACAGAAAGCTGCTCTTTTGTGCTTGA 827
Db      417 AGTCTCAGCTTTGCTTCACTTCCAGACAGAAAGCTGCTCTTTTGTGCTTGA 476
Qy      828 ACCCGAGATGAGAGATTTTGGAGCCCGTGAAGAAATGAGAGAGATGGGGACCT 887
Db      477 ACCCGAGATGAGAGATTTTGGAGCCCGTGAAGAAATGAGAGAGATGGGGACCT 536
Qy      888 TGACCTGAACGGGAGTGTGTGTGCAACAACCGCTGAGAAATCCCAACCGTCCAGA 947
Db      537 TGACCTGAACGGGAGTGTGTGTGCAACAACCGCTGAGAAATCCCAACCGTCCAGA 596
Qy      948 GGAGCTCAGAGCAGCGCTGGGAAAGCCAGACAGATGAGAGAGCGTGAATGATCAT 1007
Db      597 GGAGCTCAGAGCAGCGCTGGGAAAGCCAGACAGATGAGAGAGCGTGAATGATCAT 656

```

QY 1008 CTGCGAGACCTGCTGACGACTGCGGAGTTTGCAGCCCTGATGACAGTTTCTGCGC 1067  
DB 657 CTGCGAGACCTGCTGACGACTGCGGAGTTTGCAGCCCTGATGACAGTTTCTGCGC 716  
QY 1068 GAGCTTGTACTCGGCTGATGAGAGCGCTGCTCCTGTGTCTTACCTGCGCTGCTCCGT 1127  
DB 717 GGGCTTGTACTCGGCTGATGAGAGCGCTGCTCCTGTGTCTTACCTGCGCTGCTCCGT 776  
QY 1128 GAGCGGATCTGTAAAAACCACTCTCAACCACTGTGTGAAGCATCTCATCCAGCA 1187  
DB 777 GAGCGGATCTGTAAAAACCACTCTCAACCACTGTGTGAAGCATCTCATCCAGCA 836  
QY 1188 TCAGACAGAGTGTGAGTGAAGATGTGCAAAAGTATGATGCGCAGTAATAATCAC 1247  
DB 837 TCAGACAGAGTGTGAGTGAAGATGTGCAAAAGTATGATGCGCAGTAATAATCAC 896  
QY 1248 TCAGACAGTGTGAGTGTGAGTGAAGATGTGCAAAAGTATGATGCGCAGTAATAATCAC 1307  
DB 897 TCAGACAGTGTGAGTGTGAGTGAAGATGTGCAAAAGTATGATGCGCAGTAATAATCAC 956  
QY 1308 GAGCTTGTACTCGGCTGATGAGAGCGCTGCTCCTGTGTCTTACCTGCGCTGCTCCGT 1367  
DB 957 GAGCTTGTACTCGGCTGATGAGAGCGCTGCTCCTGTGTCTTACCTGCGCTGCTCCGT 1016  
QY 1368 CGATGCGGCGAGTGTCTGAGTACAGAGGAGGCGGCGCGCTCCCACTGCGCGCAGC 1427  
DB 1017 CGATGCGGCGAGTGTCTGAGTACAGAGGAGGCGGCGCGCTCCCACTGCGCGCAGC 1076  
QY 1428 ACCCGAGGCGAGCGAG 1487  
DB 1077 ACCCGAGGCGAGCGAG 1136  
QY 1488 CCGTGAAGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1547  
DB 1137 CCGTGAAGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1196  
QY 1548 CTGCTGCTTCCAGCGGAG 1607  
DB 1197 CTGCTGCTTCCAGCGGAG 1256  
QY 1608 CCGTGAAGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1667  
DB 1257 CCGTGAAGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1316  
QY 1668 GACCGGCTGCTAGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1702  
DB 1317 GACCGGCTGCTAGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1351

RESULT 8  
AAS25843 standard; cDNA; 1311 BP.  
ID AAS25843; XX  
AC AAS25843; XX  
XX XX  
DT 07-NOV-2001 (first entry)  
XX XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 22.  
XX XX  
KM Human; immunosuppressive; antiarthritic; as; antirheumatic; cytostatic;  
KM cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KM antibacterial; virocidic; fungicide; ophthalmological; vulnary;  
KM secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KM cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KM cerebral ischemia; angiogenesis; nervous system disorder;  
KM Alzheimer's disease; infection; ocular disorder; corneal infection;  
KM wound healing; epithelial cell proliferation; skin ageing; food additive;  
KM preservative; antiproliferative.  
XX XX  
OS Homo sapiens.  
XX XX  
PN MO20015532-A2.  
XX XX

PD 02-AUG-2001.  
XX XX  
PF 17-JAN-2001; 2001MO-US001341.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227099P.  
PR 23-AUG-2000; 2000US-0227099P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235534P.  
PR 27-SEP-2000; 2000US-0235536P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.

[illegible]

DR P-PSDB: AAU15856.

**PT** New nucleic acid molecules encoding 461 human secreted proteins for PT diagnosing, preventing, treating or ameliorating medical conditions and PT used as food additives or preservatives.

Claim 1; SEQ ID NO 22; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match	28.6%;	Score 765;	DB 4;	Length 1311;
Best Local Similarity	100.0%;	Pred. No. 2.8e-273;		
Matches 765; Conservative	0;	Mismatches	0;	Gaps 0;

QY	710	GGTCTGGGGGTGGTGGCACTCCCTTAAGGAAGTGGTCCCTCTGTGGCAAGTGAAG	769
Db	383	GGTCTGGGGGGTGGGCATCTCCCTTAAGGAAGATGGTCCCTCTGTGGCAAGTGAAG	442
QY	770	TCTCAGCTTGGCTCAGCTCTCCAGACAGAAAGA CTGCGTCCCTTTTCGTGTGAAC	829
Db	443	TCTCAGCTTGGCTCAGCTCTCCAGACAGAAAGA CTGCGTCCCTTTTCGTGTGAAC	502
QY	830	CCCAAGATCAGAGGATTTGGAGCCCGTGAAGAA GAAATGAGAGAGATGGAGCCTTG	889
Db	503	CCCAAGATCAGAGGATTTGGAGCCCGTGAAGAA AATGAGAGAGATGGAGCCTTG	562
QY	890	ACCGAAGCGGGCAGTGTGTGGTGGCAACA CCGCGTAGAAATGGCCAAACCGTCCACGAG	949
Db	563	ACCGAAGCGGGCAGTGTGTGGTGGCAACA CCGCGTAGAAATGCCAAACCGTCCACGAG	622
QY	950	ACGTCAAGACGACGGCTGGGAAGCCAGACAAGATGAGAGAGCGCTGACATGCATCT	1009
Db	623	ACGTCAAGACGACGGCTGGGAAGCCAGACAAGATGAGAGAGCGCTGACATGCATCT	682
QY	1010	GCCAAGACCTGCTGCA CGACTGGGTGAGTTTGACGCCCTGTGATGACACGTTCTGCGGG	1065
Db	683	GCCAAGACCTGCTGCA CGACTGGGTGAGTTTGACGCCCTGTGATGACACGTTCTGCGGG	742
QY	1070	CTTGTCTACTCGGGCTGGATGGAAGGCTCGTCCCTGTGTCTTACTCTGCGCGCTGTCCCGTG	1129
Db	743	CTTGTCTACTCGGGCTGGATGGAAGGCTCGTCCCTGTGTCTTACTCTGCGCGCTGTCCCGTG	802
QY	1130	AGCGGATCTGTAAAAACCAATCCTCAACAACCTCGGGAAGGATTACTCATTCACGATC	1188
Db	803	AGCGGATCTGTAAAAACCAATCCTCAACAACCTCGGGAAGGATTACTCATTCACGATC	862
QY	1190	CAGACAAGATGCGCAGTGAAGAAAGATGGAAGATGAGATGCCAGGATATAATCACTC	1245
Db	863	CAGACAAGATGCGCAGTGAAGAAAGATGGAAGATGAGATGCCAGGATATAATCACTC	922

QY 1250 AAGACATGCTGACAGCCCAAGTACAGGCGGTCTTTTCTGATGAGAGGAGTTACAGAG 1309  
 |||||  
 DB 923 AAGACATGCTGACAGCCCAAGTACAGGCGGTCTTTTCTGATGAGAGGAGTTACAGAG 982  
 |||||  
 QY 1310 ACCTGCTGAGCTGTACAGCTTGAAGTCTTCTGACATTTAGCCAGCCATACGTCG 1369  
 |||||  
 DB 983 ACCTGCTGAGCTGTACAGCTTGAAGTCTTCTGACATTTAGCCAGCCATACGTCG 1042  
 |||||  
 QY 1370 TGTGCGGCGAGTGTCTTGAATACAGAGGAGGCGGCGGCGCTCTCCCACTGCCAGCAG 1429  
 |||||  
 DB 1043 TGTGCGGCGAGTGTCTTGAATACAGAGGAGGCGGCGGCGCTCTCCCACTGCCAGCAG 1102  
 |||||  
 QY 1430 CCGAGGCGGAGCCAGGAGCCCAAGGCGGCGGCGGAGTGAACCT 1474  
 |||||  
 DB 1103 CCGAGGCGGAGCCAGGAGCCCAAGGCGGCGGCGGAGTGAACCT 1147  
 |||||  
 RESULT 9  
 ABX73184  
 ID ABX73184 standard; DNA; 1311 BP.  
 XX  
 AC ABX73184;  
 XX  
 DT 18-MAR-2003 (first entry)  
 XX  
 DE Human novel polynucleotide #12.  
 KW Human; gene; der. neural disorder; immune system disorder; renal disorder;  
 KW muscular disorder; respiratory disease; reproductive disorder;  
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KW blood related disorder; cancer; immunosuppressive; antinflammatory;  
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;  
 KW haemostatic; antiarteriosclerotic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002132753-A1.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 17-JAN-2001; 2001US-00764864.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 14-AUG-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226688P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 25-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.

PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236372P.  
 PR 29-SEP-2000; 2000US-0236388P.  
 PR 29-SEP-2000; 2000US-0236399P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239355P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 XX  
 DR WPI; 2003-147444/14.  
 DR P-PSDB; ABUS4924.  
 XX  
 PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
 PT inhibiting or preventing e.g. neural, immune system, muscular, or  
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
 PT renal disorders.  
 XX  
 PS Claim 1, SEQ ID NO 22; 402pp; English.  
 XX  
 CC The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
 CC leukemia), inflammatory diseases (e.g. septic shock, bursitis and  
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
 CC human novel polynucleotides of the invention  
 XX  
 SO Sequence 1311 BP; 306 A; 359 C; 399 G; 239 T; 0 U; 8 Other;  
 Query Match 28.6%; Score 765; DB 8; Length 1311;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-273; Indels 0; Gaps 0;  
 Matches 765; Conservative 0; Mismatches 0;  
 QY 710 GGTCTGGGGGTGGTGCATCTCCCTAAAGAAAGTGTCTCTGTGGCAAGATGAAG 769  
 |||||  
 DB 383 GGTCTGGGGGTGGTGCATCTCCCTAAAGAAAGTGTCTCTGTGGCAAGATGAAG 442  
 |||||  
 QY 770 TCTCCAGCTTGTGCTCAGCTCTCCAGAGAAAGACTGCTCTTTTGTGCTGGAAC 829  
 |||||  
 DB 443 TCTCCAGCTTGTGCTCAGCTCTCCAGAGAAAGACTGCTCTTTTGTGCTGGAAC 502  
 |||||  
 QY 830 CCGAGATCAGAGAGATTGAGCCCGTGAAGAAAGAAATGAGAGAGATGGGACTTG 889  
 |||||  
 DB 503 CCGAGATCAGAGAGATTGAGCCCGTGAAGAAAGAAATGAGAGAGATGGGACTTG 562  
 |||||  
 QY 890 ACCTGAACGGGAGTGTGTGTGCAACACCGGTGAATGCCCAACCGTCCAGAG 949  
 |||||

Db 563 ACCTGAACGGGCACTTGTGTCGACACACCGCGTAGAAATGCCCCAACCCGTCACAGAG 622  
 QY 950 ACCTCAGAGCAGCGCTGGAGACCCAGACAAAGATGAGAGAGCGTCGACATGATCATCT 1009  
 Db 623 ACCTCAGAGCAGCGCTGGAGACCCAGACAAAGATGAGAGAGCGTCGACATGATCATCT 682  
 QY 1010 GCCAGACCTGCTGACAGATGCGGTGAGTTTGACAGCCCTGTCATGACAGTTTGGCGCG 1069  
 Db 683 GCCAGACCTGCTGACAGATGCGGTGAGTTTGACAGCCCTGTCATGACAGTTTGGCGCG 742  
 QY 1070 CTGTCTACTCTGGGCTGGATGAGAGCGCTGCTCCTGATGTCCTGACCTGCGCTGCCGTG 1129  
 Db 743 CTGTCTACTCTGGGCTGGATGAGAGCGCTGCTCCTGATGTCCTGACCTGCGCTGCCGTG 802  
 QY 1130 AGCGGATCTGTAAAAACCATCTCTCAACACCTCGTGAGAGCATCTTCATCCAGCATC 1189  
 Db 803 AGCGGATCTGTAAAAACCATCTCTCAACACCTCGTGAGAGCATCTTCATCCAGCATC 862  
 QY 1190 CAGACAAAGTGGCAGTGAAGAAAGATGTGCAAGTATGATGCGAGAGATTAATAATCATC 1249  
 Db 863 CAGACAAAGTGGCAGTGAAGAAAGATGTGCAAGTATGATGCGAGAGATTAATAATCATC 922  
 QY 1250 AAGACATGCTGACAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAAGGAGTTTCAGAG 1309  
 Db 923 AAGACATGCTGACAGCCCAAGTCAAGCGGTCTTTTCTGATGAAGAAGGAGTTTCAGAG 982  
 QY 1310 ACTGCTGAGAGCTGTCAAGCGTTGAAGATGATCTCTCAACATTAGCCAGCCATTCGTG 1369  
 Db 983 ACTGCTGAGAGCTGTCAAGCGTTGAAGATGATCTCTCAACATTAGCCAGCCATTCGTG 1042  
 QY 1370 TGTGCGGCGAGTGTCTGAGTACAGAAAGGAGCGGCGAGCTCCCACTGCGCCAGCAG 1429  
 Db 1043 TGTGCGGCGAGTGTCTGAGTACAGAAAGGAGCGGCGAGCTCCCACTGCGCCAGCAG 1102  
 QY 1430 CCGAGGGCGAGCCAGGAGCCCAAGAGCCCTGCGGAGATGCACCT 1474  
 Db 1103 CCGAGGGCGAGCCAGGAGCCCAAGAGCCCTGCGGAGATGCACCT 1147  
 RESULT 10  
 ADA52592  
 ID ADA52592 standard; cDNA; 2186 BP.  
 AC  
 XX ADA52592;  
 AC  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 XX Human coding sequence, SEQ ID 160.  
 DE  
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KM Gene Therapy; human; secretory protein; membrane protein; cancer;  
 KM Inflammatory disease; osteoporosis; neurological disease; gene; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 PN EP1293569-A2.  
 PN  
 XX  
 PD 19-MAR-2003.  
 PD  
 XX  
 PF 21-MAR-2002; 2002EP-00006586.  
 PF  
 XX  
 PR 14-SEP-2001; 2001JP-00328381.  
 PR  
 XX 24-JAN-2002; 2002US-0350435P.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PA  
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y,  
 XX  
 DR WPI; 2003-395539/38.  
 DR P-PSDB; ADA54231.

XX  
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 XX  
 PS Claim 1; SEQ ID NO 160; 205bp; English.  
 CC  
 CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX  
 SQ Sequence 2186 BP; 493 A; 555 C; 539 G; 599 T; 0 U; 0 Other;  
 Query Match 23.6%; Score 633; DB 10; Length 2186;  
 Best Local Similarity 99.7%; Pred. No. 1.8e-224;  
 Matches 863; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
 QY 1796 TGACATGGAATAAATGTTGACCGAGAGCTGTGGCTCTCAAGCGGAGTGTTCGC 1855  
 Db 695 TGACATGGAATAAATGTTGACCGAGAGCTGTGGCTCTCAAGCGGAGTGTTCGC 754  
 QY 1886 TGTCTGATTAAGAGTCAAGGAGACACCGTCTGTGTTACTGCTGTGCGCTGCGCACT 1915  
 Db 755 TGTCTGATTAAGAGTCAAGGAGACACCGTCTGTGTTACTGCTGTGCGCTGCGCACT 814  
 QY 1916 TCCGTGACCTGACCTATCAGTATGCGGAGAAATTCCTGCTCCGAGTTGCGAGTGC 1975  
 Db 815 TCCGTGACCTGACCTATCAGTATGCGGAGAAATTCCTGCTCCGAGTTGCGAGTGC 874  
 QY 1976 TAACTCCCGTCTGACTGCTACTGCGGCGCTAATCTGCGCACTCAGGTAAAGTCAAC 2035  
 Db 875 TAACTCCCGTCTGACTGCTACTGCGGCGCTAATCTGCGCACTCAGGTAAAGTCAAC 934  
 QY 2036 ACGCATGAATTCATCATCTGTGTAAGAGACAGAGTTCAAAAATCAGATCCAGAG 2095  
 Db 935 ACGCATGAATTCATCATCTGTGTAAGAGACAGAGTTCAAAAATCAGATCCAGAG 994  
 QY 2096 GCCCTGAGCAGCTTTCAGACCTGAGGTGAAGAGCGGTGTTTAAATACAGAGCAA 2155  
 Db 995 GCCCTGAGCAGCTTTCAGACCTGAGGTGAAGAGCGGTGTTTAAATACAGAGCAA 1054  
 QY 2156 GCAAGTCAAGGTGTTTTCACAGCCCTGAGAGGAAAGGAGCGCAGAGTCTCCGACAGTGC 2215  
 Db 1055 ACAAGTCAAGGTGTTTTCACAGCCCTGAGAGGAAAGGAGCGCAGAGTCTCCGACAGTGC 1114  
 QY 2216 TCTGGGAGTGAATCTTCTGTGAGGCTTTTACCTCTGAGTGAAGCCCTCCCAAGAGCCC 2275  
 Db 1115 TCTGGGAGTGAATCTTCTGTGAGG--TTTACCTCTGAGTGAAGCCCTCCCAAGAGCCC 1172  
 QY 2276 GGGGGCGGAGCGCCGCTCTGAGTGAAGCGTGGGAGAGGCTGTGTGTGATCAGCAGC 2335  
 Db 1173 GGGGGCGGAGCGCCGCTCTGAGTGAAGCGTGGGAGAGGCTGTGTGTGATCAGCAGC 1232  
 QY 2336 AGAGACGAAGCTTCTGTGAACATGCGGCGCTCCGCGAGAGGGGAGTTTCTCTT 2395  
 Db 1233 AGAGACGAAGCTTCTGTGAACATGCGGCGCTCCGCGAGAGGGGAGTTTCTCTT 1292  
 QY 2396 TGTACATTTTCCGAACTACAGTTAAGACAGAGTGTGTTTTCAGAAAAGTTTCAAGGG 2455  
 Db 1293 TGTACATTTTCCGAACTACAGTTAAGACAGAGTGTGTTTTCAGAAAAGTTTCAAGGG 1352  
 QY 2456 AGAAGGCGAAGTTATCAAAAACATTGTTTCAAGAGAAAGGAGCATTAAGCTTACAGCTTA 2515  
 Db 1353 AGAAGGCGAAGTTATCAAAAACATTGTTTCAAGAGAAAGGAGCATTAAGCTTACAGCTTA 1412  
 QY 2516 CAGAGCGTACCAATATCTGCTGCTGGGAGAAACCAAGAGATTTATCTATTTTATTT 2575  
 Db 1413 CAGAGCGTACCAATATCTGCTGCTGGGAGAAACCAAGAGATTTATCTATTTTATTT 1472  
 QY 2576 TAAATAGTTTGTGCTTATCTTCTAATAGATTTAAATGTCACAACTGACCAATA 2635







PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236377P.  
 PR 29-SEP-2000; 2000US-0236378P.  
 PR 29-SEP-2000; 2000US-0236388P.  
 PR 29-SEP-2000; 2000US-0236389P.  
 PR 29-SEP-2000; 2000US-0236390P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241211P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241866P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249298P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 06-DEC-2000; 2000US-0251566P.  
 PR 06-DEC-2000; 2000US-0251688P.  
 PR 06-DEC-2000; 2000US-0251689P.  
 PR 08-DEC-2000; 2000US-0251899P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI, 2001-488783/53.  
 XX  
 XX P-PSDB; AAU16317.  
 PT New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives.

Claim 1; SEQ ID NO 483; 980pp; English.

CC The invention relates to isolated nucleic acid molecules and their  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
 CC in diagnosing a pathological condition or susceptibility to a  
 CC pathological condition. Antibodies to the proteins can also be used in  
 CC alleviating symptoms associated with the disorders and in diagnostic  
 CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays  
 CC (ELISA). Disorders which are diagnosed or treated include autoimmune  
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.  
 CC neoplasms e.g. the breast or liver, cardiovascular disorders e.g. cardiac  
 CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneu-  
 CC nervous system disorders e.g. Alzheimer's disease, infections caused by  
 CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,  
 CC and many other disorders listed in the specification. The polypeptides  
 CC can also be used to aid wound healing and epithelial cell proliferation,  
 CC to prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamin,  
 CC minerals, cofactors and other nutritional components. The present  
 CC sequence encodes a novel secreted protein of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed

Query Match 17.0%; Score 455; DB 4; Length 693;

Best Local Similarity 100.0%; Pred. No. 1.6e-158; Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1044 GCCCTGCATGACACACCTTTCGCGCGGCTTGCTCTCTCGGGGTGATGAGGCGCTGTCCTT 1103  
 DB 14 GCCCTGCATGACACACCTTTCGCGCGGCTTGCTCTCTCGGGGTGATGAGGCGCTGTCCTT 73  
 QY 1104 GTGTCTTACCTGCGCTGTGTCGCGGATCTGTAAACCAATCTTCAACAACCT 1163  
 DB 74 GTGTCTTACCTGCGCTGTGTCGCGGATCTGTAAACCAATCTTCAACAACCT 133  
 QY 1164 CGTGAAGCATACCTTATCCAGATCCAGACAGAGTCCGATGGAAGAGATGTGCAAG 1223  
 DB 134 CGTGAAGCATACCTTATCCAGATCCAGACAGAGTCCGATGGAAGAGATGTGCAAG 193  
 QY 1224 TATGATGCGAGGATTAATATCACTCAAGACATGCTGACGCCCAAGTCAAGGCGTCTT 1283  
 DB 194 TATGATGCGAGGATTAATATCACTCAAGACATGCTGACGCCCAAGTCAAGGCGTCTT 253  
 QY 1284 TTCTGATGAAGAGGAGTTCAGAGACCTGCTGAGCTGTCAAGAGTTCAGAGTGC 1343  
 DB 254 TTCTGATGAAGAGGAGTTCAGAGACCTGCTGAGCTGTCAAGAGTTCAGAGTGC 313



Db	254	TTCTGATGAAGAAAGGAGTTTCAGAGACCTGCTGAGAGCTGTACAGAGTTGACAGTGAAGTC	313
Qy	1344	CTCAGACATTAGCCAGCCATYACGTGCTGTGTGCCGCGCATGTCTCTGAGTACAGAAAGGACGC	140
Db	314	CTCAGACATTAGCCAGCCATYACGTGCTGTGTGCCGCGCATGTCTCTGAGTACAGAAAGGACGC	373
Qy	1404	GGCGCAGCCTCCCCCACTTGCCTCCAGCACACCCAGAGCCAGAGCCCCCAAGGCCCTTGGG	146
Db	374	GGCGCAGCCTCCCCCACTTGCCTCCAGCACACCCAGAGCCAGAGCCCCCAAGGCCCTTGGG	433
Qy	1464	GGATGCACCTCCAGCTCCGTCAGCCTGACGACAG	1498
Db	434	GGATGCACCTCCAGCTCCGTCAGCCTGACGACAG	468
RESULT 15			
AAH11859/C			
XX	ID	AAH11859	standard; cDNA; 518 BP.
XX	AC	AAH11859;	
XX	DT	26-JUN-2001	(first entry)
XX	DE	Human cDNA clone (3'-primer) SEQ ID NO:8654.	
XX	KM	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.	
XX	OS	Homo sapiens.	
XX	PN	EP1074617-A2.	
XX	PD	07-FEB-2001.	
XX	PF	28-JUL-2000; 2000EP-00116126.	
XX	PR	29-JUL-1999; 99JP-00248036.	
XX	PR	27-AUG-1999; 99JP-00300253.	
XX	PR	11-JAN-2000; 2000JP-00118776.	
XX	PR	02-MAY-2000; 2000JP-00183767.	
XX	PR	09-JUN-2000; 2000JP-00241899.	
XX	PA	(HELI-) HELIX RES INST.	
XX	PI	Ota T, Isegai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J,	
XX	PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX	XX	WPI; 2001-318749/34.	
XX	DR	WPI; 2001-318749/34.	
XX	PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-	
XX	PT	length cDNAs defined in the specification, and for the detection and/or	
XX	PT	diagnosis of the abnormality of the proteins encoded by the full-length	
XX	PT	cDNAs.	
XX	PS	Claim 3; SEQ ID NO 8654; 2537bp + Sequence Listing; English.	
XX	XX	The present invention describes primer sets for synthesizing 5602 full-	
XX	XX	length cDNAs defined in the specification. Where a primer set comprises:	
XX	XX	(a) an oligo-dT primer and an oligonucleotide complementary to the	
XX	XX	complementary strand of a polynucleotide which comprises one of the 5602	
XX	XX	nucleotide sequences defined in the specification; where the	
XX	XX	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
XX	XX	of an oligonucleotide comprising a sequence complementary to the	
XX	XX	complementary strand of a polynucleotide which comprises a 5'-end	
XX	XX	sequence and an oligonucleotide comprising a sequence complementary to a	
XX	XX	polynucleotide which comprises a 3'-end sequence, where the	
XX	XX	oligonucleotide comprises at least 15 nucleotides and the combination of	
XX	XX	the 5'-end sequence/3'-end sequence is selected from those defined in the	
XX	XX	specification. The primer sets can be used in antisense therapy and in	
XX	XX	gene therapy. The primers are useful for synthesizing polynucleotides,	
XX	XX	particularly full-length cDNAs. The primers are also useful for the	
XX	XX	detection and/or diagnosis of the abnormality of the proteins encoded by	
XX	XX	the full-length cDNAs. The primers allow obtaining of the full-length	
XX	XX	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and	

CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB922446 to AAB958893
CC	represent human amino acid sequences; and AAH13628 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
XX	
SQ	Sequence 518 BP; 142 A; 127 C; 108 G; 136 T; 0 U; 5 Other;
Query Match	15.9%; Score 426; DB 4; Length 518;
Best Local Similarity	99.8%; Pred. No. 8.9e-148;
Matches 476; Conservative	1; Indels 0; Gaps 0;
OY	2185 AGGGAAGGAGACGGCGGTCTCCGACAGGTGCTCTGGGGTGCATCTTCGTGGAGCTTTT 2244
DB	
OY	503 AGGGAAGGAGACGACGGGCTCCGACAGGTGCTCTGGGGTGCATCTTCGTGGAGCTTTT 444
OY	2245 ACCCTCTGATGAGAACCTCCGCCAGAGCCCCGGGGGGCGCAGCCCGCTCTCTGGTAGNC 2304
DB	443 ACCCTCTGATGAGAACCTCCGCCAGAGCCCCGGGGGGCGCAGCCCGCTCTCTGGTAGNC 384
OY	2305 GCTGGGACGGGCTCTGGTGGCATCAGACAGACAGACGAGACCCTTCTGTACATGCGGC 2364
DB	383 GCTGGGACGGGCTCTGGTGGCATCAGACAGACAGACGAGACCCTTCTGTACATGCGGC 324
OY	2365 CGTCCCGCCGAGAGGGGCAAGTTTTGCTTTTGTACATTTTCGAAACTACAGTTAAGC 2424
DB	323 CGTCCCGCCGAGAGGGGCAAGTTTTGCTTTTGTACATTTTCGAAACTACAGTTAAGC 264
OY	2425 AGAAGTCTGTTTCAGGAAAAGTTTCAAAGGAGAAGGGCAAGTTTATCAAAAAATTGTT 2484
DB	263 AAAAGTCTGTTTCAGGAAAAGTTTCAAAGGAGAAGGGCAAGTTTATCAAAAAATTGTT 204
OY	2485 TCAGAGAGAGGAGCATTAAGTTTACAGCTTACAGGACGTACACAATAATCTGCTGGTGG 2544
DB	203 TCAGAGAGAGGAGCATTAAGTTTACAGCTTACAGGACGTACACAATAATCTGCTGGTGG 144
OY	2545 AAAACCCAGAGATTTTATCTATTTTATTTTATAGTTTGGGCTATCTCTATATAA 2604
DB	143 AAAACCCAGAGATTTTATCTATTTTATTTTAAAGGTTGGCTATCTCTATATAA 84
OY	2605 GATTTAAATGTACAAACTGTAGACAAATAATTAATTTAATTTTACAATTTGAC 2661
DB	83 GATTTAAATGTACAAACTGTAGACAAATAATTAATTTAATTTTACAATTTGAC 27
RESULT 16	
ID	AAAA4336
XX	AAAA4336 standard; cDNA; 575 BP.
XX	XXXX
XX	AAAA4336;
DT	21-AUG-2000 (first entry)
XX	
DE	Human secreted expressed sequence tag SEQ ID NO:911.
XX	
KM	Human; mouse; chicken; rat; secreted expressed sequence tag; EST;
KM	expressed sequence tag; EST; probe; chemotactic; proliferative;
KM	immunomodulatory; haemostatic; chemokinetic; analgesic; haemostatic;
KM	chromonodulatory; antiinflammatory; cytotoxic; antibacterial; antifungal;
KM	antiviral; antidiabetic; antisthmatic; vulnary; antiparkinsonian;
KM	antitumor; osteopathic; neuroprotective; neotropic; antipsoriatic;
KM	cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;
KM	autoimmune disorder; multiple sclerosis; allergic condition;
KM	insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KM	lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KM	central nervous system disorder; Alzheimer's disease; stroke;
KM	Parkinson's disease; Huntington's disease; coagulation disorder;
KM	haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;
KM	infection; depression; psoriasis; ss.
OS	Homo sapiens.
XX	
FN	WO200021991-A1.
XX	

PD 20-APR-2000.  
XX  
XX 15-OCT-1999; 99MO-US024206.  
XX  
XX 15-OCT-1998; 98US-0104436P.  
XX  
PA (GEMV ) GENETICS INST INC.  
XX  
XX  
XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C,  
PI Merberg D, Treacy M, Bowman MR;  
XX  
XX WPI; 2000-317938/27.  
XX  
XX  
XX Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (ESTs), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders.  
XX  
XX  
XX Claim 1; Page 437; 803pp; English.  
XX  
XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
CC sequence tags (ESTs), isolated from human, mouse, chicken and rat tissue  
CC sources. The ESTs can have a range of activities depending on the  
CC tissues they were isolated from. The activities include: chemotactic;  
CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
CC antifungal; antiviral; antidiabetic; antiaesthetic; vulnery; antitumor;  
CC osteoprotective; neuroprotective; nootropic; antiparkinsonian; antiproliferative;  
CC cerebroprotective; anticonvulsant; and antidepressant. The ESTs can be  
CC used for gene therapy and in vaccines. The ESTs are useful as probes for  
CC the identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the ESTs. Proteins encoded by the ESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumors, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
CC in the exemplification of the present invention  
XX  
XX  
SQ Sequence 575 BP; 124 A; 168 C; 165 G; 118 T; 0 U; 0 Other;  
Query Match 10.7%; Score 286; DB 3; Length 575;  
Best Local Similarity 100.0%; Pred. No. 5,6e-96;  
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 876 AGATGGGAGCCTTGACCTGAGCGGAGTGTGGTGGCGACCAACCGGTAAATGCCCA 935  
DB 285 AGATGGGAGCCTTGACCTGAGCGGAGTGTGGTGGCGACCAACCGGTAAATGCCCA 344  
QY 936 AACCGTCCAGAGAGCGTCAAGAGCAGCGGCTGGGAAACCAAGATGAGAGAGCGT 995  
DB 345 AACCGTCCAGAGAGCGTCAAGAGCAGCGGCTGGGAAACCAAGATGAGAGAGCGT 404  
QY 996 GACATGATCATCTGCTGACAGACCTGCTGACAGACCTGCGGTGATTTGACCCCTGATGTA 1055  
DB 405 GACATGATCATCTGCTGACAGACCTGCTGACAGACCTGCGGTGATTTGACCCCTGATGTA 464  
QY 1056 CACGTTCTCGCGGAGCTTGTACTCGGGCTGATGAGAGCGCTGCTGCTTACTTACCTG 1115  
DB 465 CACGTTCTCGCGGAGCTTGTACTCGGGCTGATGAGAGCGCTGCTGCTTACTTACCTG 524  
QY 1116 CCGCTGCTCCGTTGAGAGCGGATCTGTAACCAACATCTCTCAACAC 1161  
DB 525 CCGCTGCTCCGTTGAGAGCGGATCTGTAACCAACATCTCTCAACAC 570  
RESULT 17  
AAS26143  
ID AAS26143 standard; cDNA; 357 BP.  
XX

AC AAS26143;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 322.  
XX  
XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;  
KW antibacterial; vitruclide; fungicide; ophthalmological; vulnery;  
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; skin aging; food additive;  
KW preservative; antiproliferative.  
XX  
XX  
OS Homo sapiens.  
XX  
XX WO200155322-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX  
XX 17-JAN-2001; 2001MO-US001341.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0198974P.  
XX 17-MAR-2000; 2000US-0199076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225266P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.  
XX 14-AUG-2000; 2000US-0225758P.  
XX 14-AUG-2000; 2000US-0225759P.  
XX 18-AUG-2000; 2000US-0226279P.  
XX 22-AUG-2000; 2000US-0226681P.  
XX 22-AUG-2000; 2000US-0226868P.  
XX 22-AUG-2000; 2000US-0227182P.  
XX 23-AUG-2000; 2000US-0227009P.  
XX 30-AUG-2000; 2000US-0228924P.  
XX 01-SEP-2000; 2000US-0229287P.  
XX 01-SEP-2000; 2000US-0229343P.  
XX 01-SEP-2000; 2000US-0229344P.  
XX 01-SEP-2000; 2000US-0229345P.  
XX 05-SEP-2000; 2000US-0229509P.  
XX 05-SEP-2000; 2000US-0229513P.  
XX 06-SEP-2000; 2000US-0230438P.  
XX 06-SEP-2000; 2000US-0230438P.  
XX 08-SEP-2000; 2000US-0231242P.  
XX 08-SEP-2000; 2000US-0231243P.  
XX 08-SEP-2000; 2000US-0231244P.  
XX 08-SEP-2000; 2000US-0231413P.  
XX 08-SEP-2000; 2000US-0231414P.





QY 296 CAGGTGAGTGAACCTGAGAGATACCA 322  
 |||||  
 Db 245 CAGGTGAGTGAACCTGAGAGATACCA 271  
 |||||  
 RESULT 18  
 ID ABX73484 standard; DNA; 357 BP.  
 XX  
 AC ABX73484;  
 XX  
 DT 18-MAR-2003 (first entry)  
 XX  
 DE Human novel polynucleotide #312.  
 XX  
 KM Human; gene; de; neural disorder; immune system disorder; renal disorder;  
 KM muscular disorder; respiratory disease; reproductive disorder;  
 KM gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;  
 KM hyperproliferative disorder; inflammatory disease; allergic reaction;  
 KM blood related disorder; cancer; immunosuppressive; antiinflammatory;  
 KM cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;  
 KM haemostatic; antiarteriosclerotic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002132753-A1.  
 PD 19-SEP-2002.  
 XX  
 PF 17-JAN-2001; 2001US-00764864.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-021486P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 14-AUG-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225477P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234977P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251855P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 XX  
 DR WPI; 2003-147444/14.  
 XX  
 DR P-PSDB; ABUS55224.  
 XX  
 PT New polypeptides and nucleic acids, useful in gene therapy for treating,  
 PT inhibiting or preventing e.g. neural, immune system, muscular,  
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or  
 PT renal disorders.  
 XX  
 PS Claim 1; SEQ ID NO 322; 402bp; English.  
 XX  
 CC The invention relates to human novel polypeptides and their associated  
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene  
 CC therapy for treating, inhibiting or preventing neural disorders, immune  
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis  
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.  
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left  
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage  
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and  
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and  
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood  
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial  
 CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent  
 CC human novel polynucleotides of the invention  
 XX  
 SQ Sequence 357 BP; 84 A; 87 C; 118 G; 66 T; 0 U; 2 Other;  
 Query Match 10.0%; Score 267; DB 8; Length 357;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-89;  
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 56 TCCGGGTTCCGCGCGGCGGCGGAGTGAATCCGATGAGCGGCGGAGGAGCAAGC 115  
 Db 5 TCCGGGTTCCGCGCGGCGGCGGAGTGAATCCGATGAGCGGCGGAGGAGCAAGC 64  
 QY 116 AGTCGCGCGCGCGCGAGCCTTGAGGACGCGCTCTGCGCTTGAGGCGCGGAGGAGCGAGC 175  
 Db 65 AGTCGCGCGCGCGCGAGCCTTGAGGACGCGCTCTGCGCTTGAGGCGCGGAGGAGCGAGC 124  
 QY 176 CGCAGCTCTCTTGAAGAGGCGGAGTGAATCCGATGAGCGGCGGAGGAGCGAGCTTT 235  
 Db 125 CGCAGCTCTCTTGAAGAGGCGGAGTGAATCCGATGAGCGGCGGAGGAGCGAGCTTT 184  
 QY 236 CTTTCCCGAGCAATTAATGCTCTTGAGATCACTGATGATGATGATGATGATGATGAT 295  
 Db 185 CTTTCCCGAGCAATTAATGCTCTTGAGATCACTGATGATGATGATGATGATGATGAT 244  
 QY 296 CAGGTGAGTGAACCTGAGAGATACCA 322  
 Db 245 CAGGTGAGTGAACCTGAGAGATACCA 271  
 RESULT 19  
 AAS26563  
 ID AAS26563 standard; CDNA; 354 BP.  
 XX  
 AC AAS26563;

XX 07-NOV-2001 (first entry)  
DT Human cDNA encoding a novel secreted protein, Seq ID 742.  
XX  
DE Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;  
XX cardiact; vasotropic; cerebroprotective; nootropic; neuroprotective;  
XX antibacterial; virucide; fungicide; ophthalmological; vulnery;  
XX secreted protein; rheumatoid arthritis; hyperproliferative disorder;  
XX cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
XX cerebral ischaemia; angiogenesis; nervous system disorder;  
XX Alzheimer's disease; infection; ocular disorder; corneal infection;  
XX wound healing; epithelial cell proliferation; skin ageing; food additive;  
XX preservative; antiproliferative.  
OS Homo sapiens.  
XX  
XX WO200155322-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001341.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 06-SEP-2000; 2000US-0231242P.  
PR 06-SEP-2000; 2000US-0231243P.  
PR 06-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234233P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234979P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 25-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241212P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244674P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.



	XX	17-NOV-2000; 2000US-024926AP.
	PR	17-NOV-2000; 2000US-024926SP.
	PR	17-NOV-2000; 2000US-0249287P.
	PR	17-NOV-2000; 2000US-0249299P.
	PR	17-NOV-2000; 2000US-0249300P.
	PR	01-DEC-2000; 2000US-0250160P.
	PR	01-DEC-2000; 2000US-0250391P.
	PR	05-DEC-2000; 2000US-0251030P.
	PR	05-DEC-2000; 2000US-0251988P.
	PR	05-DEC-2000; 2000US-0256719P.
	PR	06-DEC-2000; 2000US-0251479P.
	PR	08-DEC-2000; 2000US-0251856P.
	PR	08-DEC-2000; 2000US-0251868P.
	PR	08-DEC-2000; 2000US-0251869P.
	PR	08-DEC-2000; 2000US-0251989P.
	PR	08-DEC-2000; 2000US-0251990P.
	PR	11-DEC-2000; 2000US-0254037P.
	PR	05-JAN-2001; 2001US-0259678P.
	PA	(HUMA-) HUMAN GENOME SCI INC.
	XX	
	P1	Rosen CA, Barash SC, Ruben SM;
	XX	
	DR	WPI: 2001-488783/53.
	DR	P-PSDB: AAU16157.
PT	XX	New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.
PS	XX	Claim 1; SEQ ID NO 323; 980bp; English.
XX	CC	The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radiimmunoassays or enzyme linked immunosorbant assays (ELISAs). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, anglogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamin, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed
Query Match		8.9%; Score 238; DB 4; Length 354;
Best Local Similarity		100.0%; Pred. No. 3.5e-78;
Matches 238; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
OY	245	GCAATAAAGCGTCTCGAGATCATCGTAGAAATTGATGGATGAATAATCAGGTCAAGG 304
Db	51	GCATTAACCTGGTCCTCGAGATCATCGTAGAAATTGATGGATGAATAATCAGGTCAAGG 110
OY	305	TGACACTGGAAGATACCAGACCACTGAGAACAGTGATTAACAAGCTGAAGTTGTTAAGA 364
Db	111	TGACACTGGAAGATACCAGACCACTGAGAACAGTGATTAACAAGCTGAAGTTGTTAAGA 170
OY	365	AGCAGACATGCCCTTTTACAGCTGGGAGATGTCATCTTGGTGTTAACAAGAAATTAAC 424
Db	171	AGCAGACATGCCCTTTTACAGCTGGGAGATGTCATCTTGGTGTTAACAAGAAATTAAC 230





```
XX Human; gene; de; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; cancer; immunosuppressive; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotoxic; cytotoxic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
XX Homo sapiens.
XX
XX US200213753-A1.
XX
XX 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214886P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 22-AUG-2000; 2000US-0225758P.
XX 30-AUG-2000; 2000US-0226868P.
XX 01-SEP-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 08-SEP-2000; 2000US-0231413P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234979P.
XX 27-SEP-2000; 2000US-0235834P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
XX 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
XX 29-SEP-2000; 2000US-0236370P.
XX 02-OCT-2000; 2000US-0236802P.
XX 02-OCT-2000; 2000US-0237037P.
XX 02-OCT-2000; 2000US-0237038P.
XX 02-OCT-2000; 2000US-0237039P.
XX 02-OCT-2000; 2000US-0237040P.
XX 13-OCT-2000; 2000US-0239935P.
XX 20-OCT-2000; 2000US-0240960P.
XX 20-OCT-2000; 2000US-0241785P.
XX 20-OCT-2000; 2000US-0241809P.
XX 01-NOV-2000; 2000US-0244617P.
XX 17-NOV-2000; 2000US-0249299P.
XX 08-DEC-2000; 2000US-0251856P.
XX 08-DEC-2000; 2000US-0251868P.
XX 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX PA (RUBE/) RUBEN S M.
XX PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI
```

```
XX WPI; 2003-147444/14.
DR P-PSDB; ABUS55225.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 1, SEQ ID NO 323; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Epstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
XX Sequence 354 BP; 101 A; 79 C; 94 G; 77 T; 0 U; 3 Other;
XX
XX Query Match 8.9%; Score 238; DB 8; Length 354;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-78;
XX Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 245 GCATTAACCTGCTCTGTGAGATCACTGTAAGTGTGATGATGAAAATGAGTCAGG 304
XX |||||
XX 51 GCATTAACCTGCTCTGTGAGATCACTGTAAGTGTGATGATGAAAATGAGTCAGG 110
XX
XX 305 TGACACTGGAAGATACCGACAGTGGAAACGATTTAAACAAGCTGAAGTTGTAAGA 364
XX |||||
XX 111 TGACACTGGAAGATACCGACAGTGGAAACGATTTAAACAAGCTGAAGTTGTAAGA 170
XX
XX 365 ACCAGCATGCGCTTTACAGACTGGGGATGTCATCTACTTGGTGTACAGAAATGAAC 424
XX |||||
XX 171 ACCAGCATGCGCTTTACAGACTGGGGATGTCATCTACTTGGTGTACAGAAATGAAC 230
XX
XX 425 CGGAACACAACTGGCATACCTCTATGAATCTTTAAGTGAAGAAAGCATGACACA 482
XX |||||
XX 231 CGGAACACAACTGGCATACCTCTATGAATCTTTAAGTGAAGAAAGCATGACACA 288
XX
XX RESULT 23
XX ACH26450
XX ID ACH26450 standard; cDNA; 449 BP.
XX
XX ACH26450;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human adult ovary cDNA #4830.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; Biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
```

PA (DBMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW,  
XX WPI, 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
PS Claim 1; SEQ ID NO 13662; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridization). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 449 BP; 147 A; 76 C; 86 G; 135 T; 0 U; 5 Other;  
Query Match 8.1%; Score 217; DB 9; Length 449;  
Best Local Similarity 100.0%; Pred. NO. 2e-70; 0; Indels 0; Gaps 0;  
Matches 217; Conservative 0; Mismatches 0;  
QY 2445 AGTTTCAGGAGAGAGAGAGAGAGTTTATCAAAAACATTTGTCAGAGAGAGCATAG 2504  
DB 92 AGTTTCAGGAGAGAGAGAGAGAGTTTATCAAAAACATTTGTCAGAGAGAGCATAG 151  
QY 2505 TTTCAGCCTACAGAGAGAGAGAGAGAGTTTATCAAAAACATTTGTCAGAGAGAGCAT 2564  
DB 152 TTTCAGCCTACAGAGAGAGAGAGAGAGTTTATCAAAAACATTTGTCAGAGAGAGCAT 211  
QY 2565 ATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2624  
DB 212 ATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 271  
QY 2625 TAGCACAATATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2661  
DB 272 TAGCACAATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 308  
RESULT 24  
ABAI6633  
ID ABAI6633 standard; DNA; 13836 BP.  
XX  
XX ABAI6633;  
XX  
XX 23-JAN-2002 (first entry)  
XX  
XX Human nervous system related polynucleotide SEQ ID NO 8964.  
XX  
XX Human; nootropic; neuroprotective; cytosstatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antischling; antianaemic; antitachycardic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;

KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200159063-A2.  
XX  
XX 16-AUG-2001.  
XX  
XX  
XX 17-JAN-2001; 2001WO-US001334.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.

```
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX Disclosure; SEQ ID NO 8964; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB12153) and proteins
XX (AB14678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 13836 BP; 3132 A; 3517 C; 3683 G; 3504 T; 0 U; 0 Other;

Query Match 6.3%; Score 169; DB 5; Length 13836;
Best Local Similarity 100.0%; Pred. No. 6.7e-53;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 709 GGGTCGCGGGGCTGTCGTCCTCCCTTAAGGAGTGTCTCTTGCGCAAGTATGAA 768
DB 266 GGGTCGCGGGGCTGTCGTCCTCCCTTAAGGAGTGTCTCTTGCGCAAGTATGAA 325

QY 769 GTCCTCAGCTTGGCTCAGCTCTCCGAGACAGAAAGCTGCTTTTGTGTTGAA 828
DB 326 GTCCTCAGCTTGGCTCAGCTCTCCGAGACAGAAAGCTGCTTTTGTGTTGAA 385

QY 829 CCCGAGATCAGAGGATTTGAGCCCGTGAAGAAATATGAGAGAG 877
DB 386 CCCGAGATCAGAGGATTTGAGCCCGTGAAGAAATATGAGAGAG 434

RESULT 25
AAKS8374
ID AAKS8374 standard; cDNA; 825 BP.
XX
XX AAKS8374;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:3434.
DE
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX MO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001354.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
```

PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-022513P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0251799P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-483426/52.  
XX P-PSDB; AAM85593.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX

PS Claim 1; SEQ ID NO 3434; 3071bp + Sequence listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 825 BP; 216 A; 187 C; 206 G; 209 T; 0 U; 7 Other;  
Query Match 5.2%; Score 139; DB 4; Length 825;  
Best Local Similarity 99.5%; Pred. No. 1.3e-41;  
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1702 GAGCTCAACCTGGGTGACAAAGTGTCTGACGCGCTGTGAACAACAGCTACGATCA 1761  
DB 348 GAGCTCAACCTGGGTGACAAAGTGTCTGACGCGCTGTGAACAACAGCTACGATCA 407  
QY 1762 GACATCCTGAAAGATTACCTGGCAACGAGGTTTGACATGAAAAACATGTTGACCGAG 1821  
DB 408 GACATCCTGAAAGATTACCTGGCAACGAGGTTTGACATGAAAAACATGTTGACCGAG 467  
QY 1822 AGCGCTGTGGCTCTCCAGCGGAGAGTGTCTGCTGTGATTACAGAGTCAAGGAGAC 1881  
DB 468 AGCGCTGTGGCTCTCCAGCGGAGAGTGTCTGCTGTGATTACAGAGTCAAGGAGAC 527  
QY 1882 ACCGTTCTGT 1891  
DB 528 ACCGTTCTGT 537  
RESULT 26  
ABN60913/c  
ID ABN60913 standard; cDNA; 617 BP.  
AC ABN60913;  
XX  
XX 28-JUN-2002 (first entry)  
DE Human cancer related polynucleotide SEQ ID NO 880.  
XX  
XX Human; cytostatic; gene expression; gene mapping; tissue profiling;  
KW gene therapy; cancer; tumour; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200214500-A2.  
XX  
XX 21-FEB-2002.  
XX  
XX 16-AUG-2001; 2001MO-US025840.  
XX  
XX 16-AUG-2000; 2000US-0226326P.  
XX  
XX (CHIR ) CHIRON CORP.  
XX (HYSR-) HYSR INC.  
XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F,  
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;  
XX  
XX WPI; 2002-241905/29.

XX  
XX New nucleic acid for producing a polypeptide, detecting differentially  
PT expressed genes correlated with a cancerous state of a mammalian cell,  
PT and inhibiting tumor growth.  
XX  
XX Claim 1; SEQ ID NO 880; 883bp + Sequence listing; English.  
XX  
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)  
CC with cytostatic activity. The polynucleotide is used to produce a  
CC polypeptide, to detect differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell and to inhibit tumour growth. The  
CC polynucleotide is used as a probe in mapping and tissue profiling. The  
CC encoded polypeptide and antibodies to the polypeptide can also be used  
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for  
CC gene therapy. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at [ftp.wipo.int/pub/published\\_pat\\_sequences](http://ftp.wipo.int/pub/published_pat_sequences)  
XX  
SQ Sequence 617 BP; 181 A; 155 C; 159 G; 122 T; 0 U; 0 Other;  
Query Match 2.9%; Score 78; DB 6; Length 617;  
Best Local Similarity 100.0%; Pred. No. 5.1e-19;  
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1966 CCAATGAGCCGTAACATCCGCTGCTGACTGCTACTGAGGCGCGTAACCTGACAGTG 2025  
DB 385 CCAATGAGCCGTAACATCCGCTGCTGACTGCTACTGAGGCGCGTAACCTGACAGTG 326  
QY 2026 AAAGCTCACACAGCCATG 2043  
DB 325 AAAGCTCACACAGCCATG 308  
RESULT 27  
ACH14968  
ID ACH14968 standard; cDNA; 476 BP.  
XX  
XX ACH14968;  
AC ACH14968;  
XX  
XX 13-OCT-2003 (first entry)  
DE Human adult brain cDNA #2180.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRMA/) DRMANAC R T.  
XX (LABA/) LABAT I.  
XX (STAC/) STACHE-CRAIN B.  
XX (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 2180; 44bp; English.  
XX

CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome.  
CC In forensic, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC [seqdata.uspto.gov/sequence.html?DocID=20030073623](http://seqdata.uspto.gov/sequence.html?DocID=20030073623)  
CC  
XX  
SQ Sequence 476 BP; 114 A; 112 C; 135 G; 106 T; 0 U; 9 Other;  
Query Match 2.9%; Score 77; DB 9; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1.2e-18;  
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2046 ATTCAATCATCTGCTGGAACAGCAAGTTCAAAACTAAGCATCCAGAGCGCCCTGAGCA 2105  
DB 52 ATTCAATCATCTGCTGGAACAGCAAGTTCAAAACTAAGCATCCAGAGCGCCCTGAGCA 111  
QY 2106 GCTTTGAGCACTGAGG 2122  
DB 112 GCTTTGAGCACTGAGG 128  
RESULT 28  
ABN33472  
ID ABN33472 standard; DNA; 60 BP.  
XX  
AC ABN33472;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human spliced transcript detection oligonucleotide SEQ ID NO:6220.  
XX  
KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 20-JUL-2001; 2001WO-1B001903.  
XX  
PR 28-JUL-2000; 2000US-0221607P.  
PR 02-MAY-2001; 2001US-0287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Maeserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX  
PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
PT genome, useful for detecting tissue-, pathology-, and developmental-  
PT specific genes.  
XX  
PS Example 1; SEQ ID NO 6220; 47pp; English.  
XX  
CC The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
CC )transcriptome comprises messenger RNAs transcribed from multiple

CC transcription units that populate a genome. The library comprises several  
CC oligonucleotides, each capable of hybridising selectively to a set of  
CC messenger RNAs transcribed from a given transcription unit of the genome,  
CC which encodes one or more messenger RNA splice variants. The  
CC oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a particular  
CC biological or pathological state, and so allowing the detection of tissue  
CC - and pathology-specific genes such as those genes only expressed in  
CC specific tissue under a specific pathological condition; to detect  
CC developmental specific genes; and to detect RNA transcripts and splice  
CC variants of a transcriptome of a patient suffering from a particular  
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from  
CC rats, humans and mice, which are used in the exemplification of the  
CC present invention. N.B. The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
XX  
SQ Sequence 60 BP; 10 A; 12 C; 18 G; 20 T; 0 U; 0 Other;  
Query Match 2.2%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 3.4e-12;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1843 GGAGTGTTCCTGCTGCTGATTACAGACTCAAGGAGACACCGTTCTGTACTGCTGT 1902  
DB 1 GGAGTGTTCCTGCTGCTGATTACAGACTCAAGGAGACACCGTTCTGTACTGCTGT 60

RESULT 29  
AAQ22007/c  
ID AAQ22007 standard; DNA; 860 BP.  
XX  
AC AAQ22007;  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 09-JUN-1992 (first entry)  
XX  
DE Sequence encoding Cor a I allergen of hazel.  
XX  
KW Fagales; hazel; IGB; tree allergen; pollen; antibodies; T-cell response;  
KW ss.  
XX  
OS Fagales.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..483  
FT /\*tag= a  
FT /product= "Aln\_g\_1"  
FT 846..860  
FT /\*tag= b  
XX  
PN WO9202621-A.  
XX  
PD 20-FEB-1992.  
XX  
PF 08-AUG-1990; 90AT-00001668.  
XX  
PR 08-AUG-1990; 90AT-00001668.  
PR 11-APR-1991; 91US-00683831.  
XX  
PA (BIOM-) BIOMAY BIOTECHN PRO.  
XX  
PI Beltenede H, Reikhsfor A, Valenta R, Hoffmannso K, Breitenbac M;  
PI Kraft D, Rumpold H, Scheiner O;  
XX  
DR WPI; 1992-080075/10.  
DR P-PsDB; AAR21795.  
XX



PT Aln g I, Cor a I and Bet v I allergens - and DNA from alder, hazel and  
 PT birch, useful in diagnosis or therapy of allergic diseases.  
 XX  
 XX  
 PS Claim 8; Page 36; 54pp; English.  
 CC The cDNA sequence encoding the Cor a I allergen of hazel was obt'd. from  
 CC the polyA+ mRNA isolated from ripe hazel pollen. The resulting cDNA was  
 CC amplified by PCR using primers whose sequences were derived from the N-  
 CC terminal amino-acid sequence of Cor a I. The DNA fragment produced was  
 CC cloned into the pBluescript KS vector and transformed into E. coli XL1-  
 CC Blue cells. The sequence allows mammals to be tested for allergic  
 CC reactions to specific tree allergens. The derived polypeptide may be used  
 CC to challenge the mammal to elicit bronchial, conjunctival, dermal, nasal  
 CC or oral provocation. The polypeptide may be used to treat a mammal  
 CC afflicted with a pollen allergy. It is administered in an amount  
 CC sufficient to hypersensitize the mammal to Cor a I. See also AA022001-8  
 CC and AA021974. (Updated on 25-MAR-2003 to correct PR field.) (Updated on  
 CC 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS  
 CC field.)  
 XX  
 SQ Sequence 860 BP; 274 A; 150 C; 190 G; 246 T; 0 U; 0 Other;  
 Query Match 1.0%; Score 26; DB 2; Length 860;  
 Best Local Similarity 100.0%; Pred. No. 8.5;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2654 AATTGACAAAAA 2679  
 DB 754 AATTGACAAAAA 729  
 RESULT 30  
 ADF45549/C  
 ID ADF45549 standard; DNA; 1000 BP.  
 XX  
 AC ADF45549;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Arabidopsis thaliana disease stress responsive promoter SEQ ID NO:18.  
 XX  
 KW promoter; disease stress-responsive promoter; stress resistance; plant;  
 KM Alternaria brassicicola infection; infection;  
 XX disease stress resistant plant; Arabidopsis thaliana; ds.  
 OS Arabidopsis thaliana.  
 XX  
 PN JP2003284566-A.  
 XX  
 PD 07-OCT-2003.  
 XX  
 PF 29-MAR-2002; 2002JP-00095389.  
 XX  
 PR 29-MAR-2002; 2002JP-00095389.  
 XX  
 PA (RIKA) RIKAGAKU KENKUSHO.  
 XX  
 DR WPI; 2003-885429/82.  
 XX  
 PT Novel stress-responsive promoter useful for producing disease stress  
 PT resistant plants, for elucidating disease stress resistance response in  
 PT plant.  
 PS Claim 1; SEQ ID NO 18; 24pp; Japanese.  
 XX  
 CC The present invention describes a promoter (I) comprising DNA having any  
 CC one of 21 1000-nucleotide long sequences (S1, see ADF45532 to ADF45552),  
 CC DNA having nucleotides in which one or more of nucleotides of (S1) is  
 CC deleted, substituted or added and functions as disease stress-responsive  
 CC promoter (DP), and DNA which hybridizes with (S1) and functions as DP.  
 CC Also described: (1) an expression vector (II) comprising (I); (2) a  
 CC transformed organism (III) comprising (II); and (3) producing stress  
 CC resistant plant by (III). (I) responds to disease stress, where the

CC stress is caused by black sooty-mould microbe (Alternaria brassicicola)  
 CC infection. (III) is useful for producing a transgenic plant. (I) is  
 CC useful for producing a disease stress resistant plant and for elucidating  
 CC the disease stress resistance response in plant. (I) enables the  
 CC detection of stress caused by A. brassicicola infection. The present  
 CC invention represents an Arabidopsis thaliana disease stress-responsive  
 CC promoter sequence, which is used in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 1000 BP; 331 A; 162 C; 138 G; 369 T; 0 U; 0 Other;  
 Query Match 0.9%; Score 25; DB 10; Length 1000;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2655 AATTGACAAAAA 2679  
 DB 285 AATTGACAAAAA 261  
 RESULT 31  
 AAAB6229  
 ID AAAB6229 standard; cDNA; 1085 BP.  
 XX  
 AC AAAB6229;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE cDNA encoding a maize chitinase polypeptide designated ZmCh13.  
 XX  
 KW Chitinase; glucosyl hydrolase family 19; glucosyl hydrolase family 18;  
 KM pathogen control; disease resistance; molecular marker; ss.  
 XX  
 OS Zea mays.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 46..969  
 FT /\*tag= a  
 FT /product= "chitinase"  
 XX  
 PN WO200056908-A2.  
 XX  
 PD 28-SEP-2000.  
 XX  
 PF 09-MAR-2000; 2000WO-US006121.  
 XX  
 PR 24-MAR-1999; 99US-0125915P.  
 XX  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Simmons CR, Valpari N;  
 XX  
 DR WPI; 2000-628269/60.  
 XX  
 PR P-PSDB; AAB18901.  
 XX  
 PT New maize chitinase genes encoding seven chitinases of glucosyl hydrolase  
 PT family 19 are useful for enhancing disease resistance in crop plants by  
 PT modulating its expression in plants.  
 XX  
 PS Claim 1; Page 79-81; 96pp; English.  
 XX  
 CC The present sequence encodes a chitinase polypeptide. The specification  
 CC describes glucosyl hydrolase family 19 chitinases (AAB18894, AAB18896,  
 CC AAB18899, and AAB18902-05), and glucosyl hydrolase family 18 chitinases  
 CC (AAB18895, AAB18897-98 and AAB18900-01). The level of chitinase in plants  
 CC can be modulated to enhance disease resistance in crop plants and for  
 CC control of pathogens. The chitinase polynucleotides are also useful as  
 CC molecular markers for genotype in a plant, and for sequence shuffling  
 XX  
 SQ Sequence 1085 BP; 208 A; 391 C; 330 G; 156 T; 0 U; 0 Other;  
 Query Match 0.9%; Score 25; DB 3; Length 1085;  
 Best Local Similarity 100.0%; Pred. No. 19;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679  
 1061 AATTGACAAAAA 1085

RESULT 32  
 AAD39143  
 ID AAD39143 standard; cDNA, 1963 BP.  
 XX  
 AC AAD39143;  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human transporter-2 (TP-2) cDNA, 63760.  
 XX  
 KW Human; 63760 protein; transporter-2; TP-2; drug screening; gene therapy;  
 KW pharmacogenetics; central nervous system; Alzheimer's disease; epilepsy;  
 KW senile dementia; depression; cardiac related disorder; arteriosclerosis;  
 KW restenosis; myocardial infarction; cancer; cell differentiation; growth;  
 KW hormonal disorder; pituitary; thyroid; hyperthyroidism; hypothyroidism;  
 KW immune; transient hypogammaglobulinemia; immunostimulator; anorectic;  
 KW congenital X-linked infantile hypogammaglobulinemia; sugar homostasis;  
 KW anorexia; neuroprotective; antiarteriosclerotic; cardiac; cytotatic;  
 KW antidepressant; vasotropic; anticonvulsant; nootropic; metabolic;  
 KW obesity; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 67..1491  
 CDS /tag= a  
 FT /product= "Human transporter-2 (TP-2) protein"  
 FT  
 XX  
 XX MO200224911-A2.  
 XX  
 PD 28-MAR-2002.  
 XX  
 PF 19-SEP-2001; 2001WO-US029373.  
 XX  
 PF 19-SEP-2000; 2000US-0233790P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Curtis RAJ;  
 XX  
 DR WPI; 2002-479523/51.  
 DR P-PSDB; AAE24349.  
 XX  
 PT Novel human transporter 2 polypeptide referred as 63760 polypeptide which  
 PT is a member of human transporter family, useful as reagents or targets  
 PT for treating TP-2 mediated disorders such as senile dementia, depression.  
 XX  
 XX Claim 1; Fig 1; 96bp; English.  
 PS  
 XX  
 XX The invention relates to 63760 polypeptide (a human transporter family  
 CC member, referred to as transporter-2 (TP-2)) and its corresponding  
 CC nucleic acid. 63760 protein is useful for treating disorders  
 CC characterised by insufficient or excessive production of a 63760  
 CC substrate or for producing 63760 inhibitors, screening for drugs or  
 CC compounds which modulate 63760 activity. 63760 DNA is used in gene  
 CC therapy. 63760 protein, DNA and its antibody are useful for screening  
 CC assays, predictive medicine (e.g., diagnostic assays, prognostic assays,  
 CC monitoring clinical trials, and pharmacogenetics); and methods of  
 CC treatment (e.g., therapeutic and prophylactic). 63760 protein, DNA and  
 CC its antibody are useful as a novel diagnostic target and therapeutic  
 CC agent in assays applicable to treatment and diagnosis of 63760 mediated  
 CC or related disorders including central nervous system disorders such as  
 CC Alzheimer's disease, senile dementia, epilepsy, depression, etc; cardiac  
 CC related disorders such as arteriosclerosis, restenosis, myocardial  
 CC infarction, etc; cancer by modulating cellular growth, differentiation or  
 CC migration; hormonal disorders such as pituitary disorders (growth

CC disorders), thyroid disorders (hyperthyroidism, hypothyroidism), etc;  
 CC immune disorder such as transient hypogammaglobulinemia, congenital X-  
 CC linked, infantile hypogammaglobulinemia, etc; disorders associated with  
 CC sugar homostasis such as obesity, anorexia, etc. The 63760 molecules are  
 CC useful as surrogate markers for one or more disorders or disease states  
 CC or for conditions leading up to disease states, as pharmacodynamic  
 CC markers and as pharmacogenomic markers. The present sequence is human TP-  
 CC 2 cDNA  
 XX  
 SQ Sequence 1963 BP; 426 A; 505 C; 562 G; 470 T; 0 U; 0 Other;  
 XX  
 Query Match 0.9%; Score 25; DB 6; Length 1963;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679  
 1916 AATTGACAAAAA 1940

RESULT 33  
 AAD37442  
 ID AAD37442 standard; cDNA, 1963 BP.  
 XX  
 AC AAD37442;  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Human transporter TP2 cDNA #1.  
 XX  
 KW Human; ss; gene; transporter; cytotatic; anorectic; antidiabetic;  
 KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;  
 KW obesity; epilepsy; diabetes.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003143675-A1.  
 XX  
 PD 31-JUL-2003.  
 XX  
 PF 22-MAY-2002; 2002US-00154419.  
 XX  
 PF 12-MAY-2000; 2000US-0204211P.  
 PR 29-JUN-2000; 2000US-0215376P.  
 PR 31-JUL-2000; 2000US-0221769P.  
 PR 19-SEP-2000; 2000US-0233790P.  
 PR 25-SEP-2000; 2000US-0235107P.  
 PR 05-OCT-2000; 2000US-0238336P.  
 PR 14-NOV-2000; 2000US-0248364P.  
 PR 15-NOV-2000; 2000US-0248878P.  
 PR 15-DEC-2000; 2000US-0256240P.  
 PR 18-DEC-2000; 2000US-0256588P.  
 PR 21-DEC-2000; 2000US-0258028P.  
 PR 22-JAN-2001; 2001US-0263169P.  
 PR 14-MAY-2001; 2001US-00858194.  
 PR 29-JUN-2001; 2001US-00858111.  
 PR 31-JUL-2001; 2001US-00919781.  
 PR 19-SEP-2001; 2001US-00957664.  
 PR 25-SEP-2001; 2001US-00964295.  
 PR 05-OCT-2001; 2001US-00972724.  
 PR 14-NOV-2001; 2001US-000024623.  
 PR 17-DEC-2001; 2001US-00024623.  
 PR 22-JAN-2002; 2002US-00055025.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Curtis RAJ, Glucksmann MA, Meyers RE;  
 XX  
 DR WPI; 2003-851783/79.  
 DR P-PSDB; AAD37443.  
 XX  
 PT New isolated nucleic acid, useful for preparing a composition for  
 PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy



Db	1268 AATTGACAAAAA	AAAAAAAAAAAA	1244
RESULT 36			
ID	ADK67852	standard; cDNA, 2424 BP.	
XX	ADK67852		
AC	ADK67852;		
XX			
DT	06-MAY-2004	(first entry)	
DE	Phosphadenosine-phosphosulfate synthetase, modifier of AXIN pathway.		
KM	Human; phosphadenosine-phosphosulfate synthetase; PAPSS; enzyme; AXIN;		
KW	cytostatic; gene therapy; gene; ss.		
OS	Homo sapiens.		
FN	WO200401309-A2.		
PD	12-FEB-2004.		
PF	06-AUG-2003; 2003WO-US024562.		
PR	06-AUG-2002; 2002US-0401534P.		
PA	(EXEL-) EXELIXIS INC.		
PI	Gendreau SB, Dora EG, Lickteig K;		
PT	WPI. 2004-157123/15.		
DR	GENBANK; 14602765.		
XX			
PT	Identifying a candidate AXIN pathway-modulating agent, useful in		
PT	diagnosing and treating breast, colon, head and neck, kidney, lung,		
PT	ovary, prostate or skin cancer, by providing an assay system comprising a		
PT	PAPSS polypeptide or nucleic acid.		
XX			
PS	Disclosure; SEQ ID NO 8; 64pp; English.		
CC	The present sequence is that of a human phosphadenosine-phosphosulfate		
CC	synthetase (PAPSS) polynucleotide. A reduction of function pry-1 (axin)		
CC	mutant in Caenorhabditis elegans was used to identify the TIGD10.1 gene		
CC	as a modifier of the AXIN pathway. Human orthologues, termed PAPSS, of		
CC	this gene were then identified. These are attractive drug targets for the		
CC	treatment of pathologies associated with a defective AXIN signalling		
CC	pathway, such as cancer. In vitro and in vivo methods of assessing PAPSS		
CC	function are provided. Modulation of the PAPSS or its binding partner can		
CC	be used to examine the association of the AXIN pathway and its members in		
CC	healthy and disease conditions and to develop diagnostic and therapeutic		
CC	modalities for AXIN related pathologies. PAPSS-modulating agents that act		
CC	by inhibiting or enhancing PAPSS expression, directly or indirectly, e.g.		
CC	by affecting a PAPSS function such, can be identified using methods of		
CC	the invention. Preferred modulating agents include antisense oligomers		
CC	and phosphorothioate morpholino oligomers.		
XX			
SQ	Sequence 2424 BP; 661 A; 565 C; 596 G; 602 T; 0 U; 0 Other;		
Query Match	0.9%; Score 25; DB 12; Length 2424;		
Best Local Similarity	100.0%; Pred. No. 17;		
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
OY	2655 AATTGACAAAAA	AAAAAAAAAA	2679
DB	2394 AATTGACAAAAA	AAAAAAAAAA	2418
RESULT 37			
ID	ADA68659		
NC	ADA68659 standard; DNA; 95769 BP.		
XX	ADA68659;		
XX			

DT	20-NOV-2003	(first entry)
XX		
DE	Arabidopsis thaliana gene, SEQ ID 1979.	
XX		
KW	Plant; bacterial infection; fungal infection; viral infection; ds.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	WO200300898-A1.	
XX		
XX		
PD	03-JAN-2003.	
XX		
PF	22-JUN-2001; 2001WO-IB001105.	
XX		
PR	22-JUN-2001; 2001WO-IB001105.	
XX		
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.	
XX		
XI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
PI	Katagiri F, Quan S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;	
XX		
DR	WPI, 2003-175290/17.	
PT		
PT	Identifying at least one gene involved in plant resistance or response to	
PT	pathogenic infection for conferring resistance or tolerance to a plant to	
PT	bacterial, fungal or viral infection by determining or detecting plant	
PT	gene expression.	
XX		
PS		
XX	Disclosure; SEQ ID NO 1979; 899pp; English.	
CC		
CC	The present invention relates to a method (M1) for identifying genes	
CC	involved in plant resistance or response to pathogenic infection. M1	
CC	comprises identifying a gene whose expression is significantly altered in	
CC	the incompatible interaction of plant gene expression relative to	
CC	expression of the gene in an uninfected plant, in a mutant plant that	
CC	does not express a gene associated with response to pathogenic infection,	
CC	or in a corresponding incompatible or compatible interaction. (M1) is	
CC	useful for conferring resistance to resistance or tolerance to a plant to	
CC	bacterial, fungal or viral infection. The present sequence was used to	
CC	illustrate the invention.	
SQ		
Sequence	95769 BP; 30958 A; 16961 C; 16682 G; 31168 T; 0 U; 0 Other;	
Query Match	0.9%; Score 25; DB 8; Length 95769;	
Best Local Similarity	100.0%; Pred. No. 9.4;	
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	2655 AATTGACAAAAAAAAAAAAAAAAAAAAA 2679       5271 AATTGACAAAAAAAAAAAAAAAAAAAAA 5295	
DB		
RESULT 38		
ABL86644		
ID	ABL86644 standard; cDNA; 199 BP.	
AC		
XX	ABL86644;	
DT	17-MAY-2002 (first entry)	
XX		
DE	Human ovarian cancer related cDNA clone SEQ ID NO:9622.	
XX		
KW	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200192581-A2.	
XX		
PD	06-DEC-2001.	
XX		
PE	29-MAY-2001; 2001WO-US017756.	
XX		
PR	26-MAY-2000; 2000US-0207484P.	

XX (CORI-) CORIXA CORP.  
XX  
XX Algate PA, Harlocker SL, Jones R;  
XX  
XX MPI; 2002-122075/16.  
XX  
XX Composition for therapy and diagnosis of ovarian cancer comprising  
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
PT polypeptide, antibody specific to polypeptide or T cell expressing  
PT polypeptide.  
XX  
XX Claim 1; SEQ ID NO 9622; 489pp; English.  
XX  
XX The present invention describes a composition (I) comprising: carriers  
CC and immunostimulants; and a polypeptide (II) of a ovarian tumor  
CC polypeptide encoded by a polynucleotide (III) having a CDNA sequence (S1)  
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
CC or antigen presenting cells that express (II), (I) has cytostatic  
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for  
CC detecting ovarian cancer in a patient's biological sample preferably  
CC serum or ovarian tissue. The method comprises contacting a biological  
CC sample from a patient with (IV), detecting the amount of polynucleotide  
CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
CC value and thereby detecting ovarian cancer in the patient, where the  
CC amount of polynucleotide hybridising to (IV) is detected preferably by  
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
CC useful for stimulating and/or expanding T cells specific for an ovarian  
CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
CC useful in design and preparation of ribozyme molecules for inhibiting  
CC expression of the tumour polypeptides and proteins in tumour cells; and  
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
CC library using well known techniques  
XX  
SQ Sequence 199 BP; 81 A; 36 C; 49 G; 33 T; 0 U; 0 Other;  
XX  
Query Match 0.94; Score 24; DB 6; Length 199;  
Best Local Similarity, 100.0%; Pred. No. 58;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2656 ATTGACMAAAAAAAAAAAAAA 2679  
DB 160 ATTGACMAAAAAAAAAAAAAA 183  
RESULT 39  
AAS29141  
ID AAS29141 standard; CDNA; 222 BP.  
XX  
XX AAS29141;  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX  
DE CDNA encoding for human DNA-binding protein #112.  
XX  
XX  
XX Human; DNA-binding protein; histone; chromo domain protein;  
XX chromatin organisation modifier; Y-box binding protein; DNA organisation;  
XX gene transcription; malignant disease; autoimmune disorder;  
XX rheumatic disease; genetic abnormality; infectious disease;  
XX neurological disorder; gene therapy; immunomodulatory; anti-HIV;  
XX anti rheumatic; anti microbial; cytostatic; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200155162-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001MO-US001305.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX  
XX 04-FEB-2000; 2000US-0180628P.  
XX

PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220863P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.





PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226661P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234977P.  
 PR 25-SEP-2000; 2000US-0234988P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241211P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251899P.  
 PR 08-DEC-2000; 2000US-0251900P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Rosen CA, Barash SC, Ruben SM,  
 XX MPI; 2001-465557/50.  
 DR P-PSDB; AAU18193.  
 XX  
 PT Nucleic acid molecules encoding human secreted chromosomal binding  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers.  
 XX  
 XX  
 XX Claim 4; SEQ ID NO 50; 561pp; English.  
 XX  
 CC The present invention relates to the isolation of novel DNA-binding  
 CC proteins (AAU1814-AAU1821), and cDNA and genomic sequences encoding for  
 CC these proteins. DNA-binding proteins such as histones, chromo (chromatin  
 CC organisation modifier) domain proteins, and Y-box binding proteins may  
 CC contribute to diseases resulting from aberrant DNA organisation and/or  
 CC gene transcription. The sequences of the invention are useful in  
 CC screening assays to identify antagonists and/or agonists that may enhance  
 CC or block activities mediated by DNA-binding proteins. Blockers of DNA-  
 CC binding proteins may be useful in treating disorders such as malignant  
 CC diseases (e.g. cancer), autoimmune disorders (e.g. diabetes mellitus),  
 CC rheumatic diseases (e.g. rheumatoid arthritis), genetic abnormalities  
 CC (e.g. cystic fibrosis), infectious diseases (e.g. HIV) and neurological

CC disorders (e.g. Alzheimer's disease). The polynucleotide sequences of the  
 CC invention may also be used in gene therapy. AAS29030-AAS29157 represent  
 CC cDNA sequences encoding for novel DNA-binding proteins. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 222 BP; 112 A; 26 C; 25 G; 57 T; 0 U; 2 Other;

Query Match 0.9%; Score 24; DB 5; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679

DB 145 ATTGACAAAAA 168

RESULT 41

AB568281  
 ID ABS68281 standard; cDNA; 222 BP.

AC ABS68281;

XX 18-NOV-2002 (first entry)

DE cDNA encoding human DNA-binding protein #112.

KW Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;  
 KW severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;  
 KW diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;  
 KW graft-versus-host disease; blood-related disorder; atherosclerosis;  
 KW hyperproliferative disorder; cancer; renal disorder; arrhythmia;  
 KW acute glomerulonephritis; cardiovascular disorder; respiratory disorder;  
 KW Goodpasture's syndrome; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; endocrine disorder; Addison's disease; gene;  
 KW reproductive system disorder; endometriosis; infectious disease;  
 KW viral infection; bacterial infection; fungal infection; vaccine;  
 KW gastrointestinal disorder; multiple sclerosis; gene therapy; ss.

OS Homo sapiens.

XX US2002102638-A1.

XX 01-AUG-2002.

PF 17-JAN-2001; 2001US-00764846.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0218390P.

PR 26-JUL-2000; 2000US-0220863P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 22-AUG-2000; 2000US-0226688P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 08-SEP-2000; 2000US-0231413P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234937P.

PR 27-SEP-2000; 2000US-0235834P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0244617P.

PR 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

XX (RUBE/) RUBEN S M.

XX (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI, 2002-690611/74.

XX P-PSDB; ABG92686.

DR Novel DNA-binding protein useful for diagnosis, prognosis, prevention and

PT treatment of immune, hyperproliferative, respiratory, cardiovascular,

PT reproductive, endocrine, gastrointestinal and neurological disorders.

XX Claim 1; SEQ ID NO 122; 225bp; English.

XX The present invention relates to a new DNA-binding protein. The invention

XX is useful in treating, preventing, diagnosing and/or prognosing

XX immunodeficiencies (e.g. B cell immunodeficiencies, severe combined

XX immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple

XX sclerosis, diabetes mellitus), allergic reactions and conditions (e.g.

XX asthma), inflammatory conditions, graft-versus-host disease, blood-

XX related disorders (thrombosis, atherosclerosis), hyperproliferative

XX disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis),

XX cardiovascular disorders (e.g. arrhythmia), respiratory disorders

XX (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's

XX disease, Parkinson's disease), endocrine disorders (e.g. Addison's

XX disease), reproductive system disorders (e.g. endometriosis), infectious

XX diseases (e.g. viral, bacterial or fungal infections) and

XX gastrointestinal disorders (e.g. Crohn's disease). The invention is also

XX useful to stimulate neuronal growth and treat, prevent, and/or diagnose

XX neuronal damage which occurs in certain neuronal disorders or neuro-

XX degenerative conditions. The present nucleic acid sequence encodes a

XX human DNA-binding protein of the invention. Note: The sequence data for

XX this patent did not form part of the printed specification, but was

XX obtained in electronic format directly from USPTO at

XX http.seqdata.uspto.gov/sequence

XX Sequence 222 BP; 112 A; 26 C; 25 G; 57 T; 0 U; 2 Other;

Query Match 0.9%; Score 24; DB 6; Length 222;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679

DB 145 ATTGACAAAAA 168

RESULT 42  
 ABS68209  
 ID ABS68209 standard; cDNA; 222 BP.  
 AC ABS68209;  
 DT 18-NOV-2002 (first entry)  
 XX  
 DE cDNA encoding human DNA-binding protein #40.  
 XX  
 KW Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;  
 KW severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;  
 KW diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;  
 KW graft-versus-host disease; blood-related disorder; atherosclerosis;  
 KW hyperproliferative disorder; cancer; renal disorder; arrhythmia;  
 KW acute glomerulonephritis; cardiovascular disorder; respiratory disorder;  
 KW Goodpasture's disease; endocrine disorder; Alzheimer's disease;  
 KW Parkinson's disease; endocrine disorder; Addison's disease; gene;  
 KW reproductive system disorder; endometriosis; infectious disease;  
 KW viral infection; bacterial infection; fungal infection; vaccine;  
 KW gastrointestinal disorder; multiple sclerosis; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002102638-A1.  
 XX  
 PD 01-ANG-2002.  
 XX  
 PF 17-JAN-2001; 2001US-00764846.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 14-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226868P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234977P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.

PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX  
 PA (ROSE/) ROSEN C A.  
 PA (RUBEN/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 XX  
 DR WPI; 2002-690611/74.  
 DR P-PDB; ABG92614.  
 XX  
 PT Novel DNA-binding protein useful for diagnosis, prognosis, prevention and  
 PT treatment of immune, hyperproliferative, respiratory, cardiovascular,  
 PT reproductive, endocrine, gastrointestinal and neurological disorders.  
 XX  
 PS Claim 1; SEQ ID NO 50; 225bp; English.  
 XX  
 CC The present invention relates to a new DNA-binding protein. The invention  
 CC is useful in treating, preventing, diagnosing and/or prognosing  
 CC immunodeficiencies (e.g. B cell immunodeficiencies, severe combined  
 CC immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple  
 CC sclerosis, diabetes mellitus), allergic reactions and conditions (e.g.  
 CC asthma), inflammatory conditions, graft-versus-host disease, blood-  
 CC related disorders (thrombosis, atherosclerosis), hyperproliferative  
 CC disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis),  
 CC cardiovascular disorders (e.g. arrhythmia), respiratory disorders  
 CC (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's  
 CC disease, Parkinson's disease), endocrine disorders (e.g. Addison's  
 CC disease), reproductive system disorders (e.g. endometriosis), infectious  
 CC diseases (e.g. viral, bacterial or fungal infections) and  
 CC gastrointestinal disorders (e.g. Crohn's disease). The invention is also  
 CC useful to stimulate neuronal growth and treat, prevent, and/or diagnose  
 CC neuronal damage which occurs in certain neuronal disorders or neuro-  
 CC degenerative conditions. The present nucleic acid sequence encodes a  
 CC human DNA-binding protein of the invention. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC http://seqdata.uspto.gov/sequence  
 XX  
 SQ Sequence 222 BP; 112 A; 26 C; 25 G; 57 T; 0 U; 2 Other;  
 XX  
 QY Query Match 0.9%; Score 24; DB 6; Length 222;  
 QY Best Local Similarity 100.0%; Pred. No. 57;  
 Db Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679  
 Db 145 ATTGACAAAAAAAAAAAAAAAAAAAA 168  
 XX  
 AC ADC25203;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human cDNA from extracellular matrix gene 40.  
 XX  
 KW Extracellular matrix protein; cytosolic; antibacterial; virucide;  
 KW neuroprotective; gynaecological; gastrointestinal-gen; candidant;  
 KW cardiovascular-gen; nephrotropic; antiinflammatory; muscular-gen;  
 KW respiratory-gen; immunosuppressive; cerebroprotective; vasotropic;  
 KW neurotropic; antiallergic; cancer; bacterial infection; viral infection;  
 KW neural disorder; immune system disorder; blood disorder;

KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW inflammatory disorder; proliferative disorder; human; gene therapy; ss;  
KW gene.  
XX Homo sapiens.  
XX US2003049650-A1.  
XX PD 13-MAR-2003.  
XX PD 07-MAR-2002; 2002US-00091483.  
XX PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 29-SEP-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249219P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251969P.

PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764846.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX WPI, 2003-605749/57.  
DR P-PSDB; ADC25331.  
XX  
XX New DNA-binding proteins and gene encoding them, useful for diagnosing,  
PT creating and/or preventing e.g. neurological, inflammatory, infectious,  
PT cardiovascular, autoimmune, respiratory, neoplastic or digestive  
PT diseases.  
XX  
XX Claim 1; SEQ ID NO 50; 226bp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (cDNA)  
CC encoding a human extracellular matrix protein, representing one of 161  
CC novel genes. Also included are recombinant vectors, host cells  
CC (expressing the protein), the extracellular matrix proteins (including  
CC their fragments, epitopes and homologues), an isolated antibody that  
CC binds specifically to the protein, diagnosing a pathological condition or  
CC susceptibility to a pathological condition (comprising determining the  
CC presence or absence of a mutation in the nucleic acid and diagnosing a  
CC condition based on the presence or absence of the mutation), diagnosing a  
CC pathological condition or susceptibility to a pathological condition  
CC (comprising determining the presence or amount of expression of the  
CC protein in a biological sample and diagnosing a condition based on the  
CC presence or amount of expression of the protein), preventing, treating or  
CC ameliorating a medical condition by administering the nucleic acid or  
CC protein to a mammalian subject, identifying a binding partner to the  
CC protein, the gene corresponding to the cDNA sequence, and identifying an  
CC activity in a biological assay (comprising expressing the nucleic acid in  
CC a cell, isolating the supernatant, detecting an activity in a biological  
CC assay and identifying the protein in the supernatant having the  
CC activity). The nucleic acids and proteins display the following  
CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,  
CC

Query Match 0.9%; Score 24; DB 10; Length 222;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2656 ATTGACAAAAA 2679  
Db 145 ATTGACAAAAA 168

RESULT 44  
ADC25275  
ID ADC25275 standard; cDNA; 222 BP.  
XX  
XX  
AC ADC25275;  
XX  
XX 18-DEC-2003 (first entry)  
DT  
XX  
XX Human cDNA from extracellular matrix gene 112.  
DE  
XX  
XX Extracellular matrix protein; cytoskeletal; antibacterial; virucide;  
XX neuroprotective; gynaecological; gastrointestinal; cardiac;  
XX cardiovascular; nephrotropic; anti-inflammatory; muscular; gen;  
XX respiratory; immunosuppressive; cerebroprotective; vasotrophic;  
XX neurotropic; antitumor; cancer; bacterial infection; viral infection;  
XX neural disorder; immune system disorder; blood disorder;  
XX muscular disorder; reproductive disorder; gastrointestinal disorder;  
XX pulmonary disorder; cardiovascular disorder; renal disorder;  
XX inflammatory disorder; proliferative disorder; human; gene therapy; ss;  
XX gene.  
OS Homo sapiens.  
XX

PN US2003049650-A1.  
XX  
XX 13-MAR-2003  
XX  
XX 07-MAR-2002; 2002US-00091483.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225679P.  
PR 14-AUG-2000; 2000US-0225688P.  
PR 14-AUG-2000; 2000US-0225709P.  
PR 14-AUG-2000; 2000US-0225709P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226882P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228987P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229511P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241866P.  
 PR 20-OCT-2000; 2000US-0241877P.  
 PR 20-OCT-2000; 2000US-0241888P.  
 PR 20-OCT-2000; 2000US-0241899P.  
 PR 01-NOV-2000; 2000US-0241826P.  
 PR 08-NOV-2000; 2000US-024617P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249247P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 01-DEC-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 05-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251866P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2000US-0259678P.  
 PR 17-JAN-2001; 2000US-00764846.  
 XX  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX

PI Rosen CA, Ruben SM, Barash SC;  
 XX  
 XX WPI; 2003-605749/57.  
 DR P-PSDB; ADC25403.  
 XX  
 PT New DNA-binding proteins and gene encoding them, useful for diagnosing,  
 PT treating and/or preventing e.g. neurological, inflammatory, infectious,  
 PT cardiovascular, autoimmune, respiratory, neoplastic or digestive  
 diseases.  
 XX  
 PS Claim 1; SEQ ID NO 122; 226bp; English.  
 CC  
 CC The invention relates to an isolated nucleic acid molecule (cDNA)  
 CC encoding a human extracellular matrix protein, representing one of 161  
 CC novel genes. Also included are recombinant vectors, host cells  
 CC (expressing the protein), the extracellular matrix proteins (including  
 CC their fragments, epitopes and homologues), an isolated antibody that  
 CC binds specifically to the protein, diagnosing a pathological condition or  
 CC susceptibility to a pathological condition (comprising determining the  
 CC presence or absence of a mutation in the nucleic acid and diagnosing a  
 CC condition based on the presence or absence of the mutation), diagnosing a  
 CC pathological condition or susceptibility to a pathological condition  
 CC (comprising determining the presence or amount of expression of the  
 CC protein in a biological sample and diagnosing a condition based on the  
 CC presence or amount of expression of the protein), preventing, treating or  
 CC ameliorating a medical condition by administering the nucleic acid or  
 CC protein to a mammalian subject, identifying a binding partner to the  
 CC protein, the gene corresponding to the cDNA sequence, and identifying an  
 CC activity in a biological assay (comprising expressing the nucleic acid in  
 CC a cell, isolating the supernatant, detecting an activity in a biological  
 CC assay and identifying the protein in the supernatant having the  
 CC activity). The nucleic acids and proteins display the following  
 CC activities Cytostatic, antibacterial, Virolytic, Neuroprotective,  
 CC  
 Query Match 0.9%; Score 24; DB 10; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679  
 DB 145 ATTGACAAAAAAAAAAAAAAAAAAAA 168  
 RESULT 45  
 ABV20036 standard; cDNA; 291 BP.  
 XX  
 AC ABV20036;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 20027.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX



PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 3266; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (1) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (1) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 291 BP; 148 A; 32 C; 39 G; 69 T; 0 U; 3 Other;

Query Match 0.9%; Score 24; DB 5; Length 291;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACMAAAAAAAAAAAAAAAAA 2679  
Db 199 ATTGACMAAAAAAAAAAAAAAAAA 222

RESULT 46  
ABV49795  
ID ABV49795 standard; cDNA; 413 BP.  
XX  
AC ABV49795;  
XX  
DT 17-SEP-2002. (first entry)  
XX  
DE Human prostate expression marker cDNA 49786.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 9711; 11750pp; English.

XX  
CC The invention relates to an isolated nucleic acid molecule (1) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (1) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 413 BP; 184 A; 96 C; 49 G; 72 T; 0 U; 12 Other;

Query Match 0.9%; Score 24; DB 5; Length 413;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACMAAAAAAAAAAAAAAAAA 2679  
Db 224 ATTGACMAAAAAAAAAAAAAAAAA 247

RESULT 47  
ACH23460  
ID ACH23460 standard; cDNA; 474 BP.  
XX  
AC ACH23460;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
DE Human adult ovary cDNA #1840.  
XX  
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2003073623-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 30-JUL-2001; 2001US-00918995.  
XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
DR WPI; 2003-615964/58.  
XX  
PT New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
PS Claim 1; SEQ ID NO 10672; 44pp; English.

CC The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations

CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC [seqdata.uspto.gov/sequence.html?DocID=20030073623](http://seqdata.uspto.gov/sequence.html?DocID=20030073623)  
CC  
XX  
SQ Sequence 474 BP; 169 A; 71 C; 78 G; 154 T; 0 U; 2 Other;  
  
Query Match 0.9%; Score 24; DB 9; Length 474;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 449 ATTGACAAAAAAAAAAAAAAAAAAAA 472  
  
RESULT 48  
ABV60912/c  
ID ABV60912 standard; cDNA; 525 BP.  
XX  
AC ABV60912;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 60903.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 11581; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (1) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (1) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker

XX  
SQ Sequence 525 BP; 112 A; 121 C; 93 G; 199 T; 0 U; 0 Other;  
  
Query Match 0.9%; Score 24; DB 5; Length 525;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 35 ATTGACAAAAAAAAAAAAAAAAAAAA 12  
  
RESULT 49  
ACC60485  
ID ACC60485 standard; cDNA; 532 BP.  
XX  
AC ACC60485;  
XX  
DT 23-JUN-2003 (first entry)  
XX  
DE Rice leaf EST, SEQ ID NO:36.  
XX  
KW Rice leaf EST; expressed sequence tag; biochip; DNA array; cloning;  
KW detection; heterosis; hybrid vigour; transgene detection;  
KW herbicide screening; pesticide screening; disease diagnosis;  
KW medical research; agriculture; ss.  
XX  
OS Oryza sativa.  
XX  
PN CN1364932-A.  
XX  
PD 21-AUG-2002.  
XX  
PF 31-OCT-2001; 2001CN-00137662.  
XX  
PR 31-OCT-2001; 2001CN-00137662.  
XX  
PA (UYZH-) UNIV ZHEJIANG.  
XX  
PI Li D, Dong H;  
XX  
DR WPI; 2003-157852/16.  
XX  
PT Rice leaf expression sequence labels and constituted biochip.  
XX  
PS Claim 1; Page 23 (Disclosure); 29pp; Chinese.  
XX  
CC The invention relates to 50 novel ESTs (expressed sequence tags; ACC60450  
CC -ACC60499) obtained from a rice leaf cDNA library. The invention also  
CC relates to a biochip comprising these ESTs. The biochip of the invention  
CC can be used in a variety of agricultural applications. It can be used in  
CC the cloning of genes which confer useful characteristics in crop species,  
CC to predict hybrid vigour (heterosis) at an early stage, in the detection  
CC of transgenic agricultural produce, in screening for novel herbicides and  
CC pesticides and in disease diagnosis. The biochip may also be used in  
CC medical research. The present sequence represents a rice leaf EST of the  
CC invention  
XX  
SQ Sequence 532 BP; 140 A; 108 C; 118 G; 166 T; 0 U; 0 Other;  
  
Query Match 0.9%; Score 24; DB 8; Length 532;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 507 ATTGACAAAAAAAAAAAAAAAAAAAA 530  
  
RESULT 50  
ACF56929  
ID ACF56929 standard; cDNA; 532 BP.  
XX

AC ACF56929;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Rice leaf EST, SEQ ID NO:36.  
XX  
KM Rice leaf EST; expressed sequence tag; biochip; DNA array; cloning;  
KM detection; heterosis; hybrid vigour; transgene detection;  
KM herbicide screening; pesticide screening; disease diagnosis;  
KM safety assessment; agriculture; plant; ss.  
XX  
OS Oryza sativa.  
XX  
PN CN1366077-A.  
XX  
PD 28-AUG-2002.  
XX  
PF 31-OCT-2001; 2001CN-00137666.  
XX  
PR 31-OCT-2001; 2001CN-00137666.  
XX  
PA (UYZH-) UNTV ZHEJIANG.  
XX  
PI Dong H, L1 D;  
XX  
DR WPI; 2003-382549/37.  
XX  
PT Rice leaf specific expression sequence label and its biochip prepared by  
XX it.  
XX  
PS Claim 1; Page 23 (Disclosure); 0pp; Chinese.  
XX  
CC The invention relates to 50 novel ESTs (expressed sequence tags; ACF56929  
CC -ACF56943) obtained from a rice leaf cDNA library. The invention also  
CC relates to a biochip comprising these ESTs. The biochip of the invention  
CC can be used in a variety of agricultural applications. It can be used in  
CC the cloning of genes which confer useful characteristics in crop species,  
CC to predict hybrid vigour (heterosis) at an early stage, in the detection  
CC of transgenic agricultural produce, in screening for novel herbicides and  
CC pesticides and in disease diagnosis. The biochip may also be used to  
CC assess the safety of transgenic agricultural products. The present  
CC sequence represents a rice leaf EST of the invention  
XX  
SQ Sequence 532 BP; 140 A; 108 C; 118 G; 166 T; 0 U; 0 Other;  
Query Match 0.9%; Score 24; DB 10; Length 532;  
Best Local Similarity 100.0%; Pred.No. 50;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 507 ATTGACAAAAAAAAAAAAAAAAAAAA 530

Search completed: January 15, 2005, 07:06:48  
Job time : 1324 secs

***This Page Blank (uspto)***

















974	23	0.9	191	6	CF227874	PCAXM0005
975	23	0.9	193	2	AM023621	dfe57h06.y
976	23	0.9	194	6	CG887261	RS4408.Ba
977	23	0.9	195	6	CB278969	ru39a09.y
978	23	0.9	195	6	CF331404	NAC--07-
979	23	0.9	198	1	A1024310	ov67E04.x
980	23	0.9	198	1	A1223559	qk29h07.x
981	23	0.9	198	1	AV741707	AV741707
982	23	0.9	204	5	BM971317	UI-CF-BE1
983	23	0.9	206	1	A1613569	ct80d07.x
984	23	0.9	207	1	A1336651	ct94g03.x
985	23	0.9	208	1	AV916959	AV916959
986	23	0.9	209	2	BE047737	tz41e01.y
987	23	0.9	210	4	BM481210	533107.MA
988	23	0.9	210	6	CF203048	RR890915N
989	23	0.9	212	4	B1535579	399194.MA
990	23	0.9	213	4	BI069737	C006P75U
991	23	0.9	213	6	CB046117	NISC_gf02
992	23	0.9	214	2	AM102133	sd83f02.y
993	23	0.9	214	2	BE115841	UI-R-B51-
994	23	0.9	214	2	BE396639	UI-M-B21-
995	23	0.9	214	9	AL937996	Arabidops
996	23	0.9	215	4	BG630048	CC-estfcl
997	23	0.9	216	1	A1208643	q934c04.x
998	23	0.9	216	2	AW827278	xn11a10.y
999	23	0.9	218	2	BF468182	UI-M-CCO-
1000	23	0.9	218	4	BM526225	sal38e10.

## ALIGNMENTS

RESULT 1  
 HSM802304  
 LOCUS HSM802304 2297 bp mRNA linear HTC 04-AUG-2004  
 DEFINITION Homo sapiens mRNA; cDNA DKFp434N2420 (from clone DKFp434N2420).  
 ACCESSION AL137561  
 VERSION AL137561.1 GI:6808265  
 KEYWORDS HTC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2297)  
 Othenwaelder, B., Obermaier, B., Deutschenbauer, S., Schaipp, A., Mewes, H.W., Weill, B., Amid, C., Oesanger, A., Fobo, G., Han, M. and Wiemann, S.

COMMENT  
 The German cDNA Consortium  
 Direct Submission  
 Submitted (03-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 Sequencing consortium of the German Genome Project.  
 This clone (DKFp434N2420) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
 Please contact RZPD for ordering:  
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFp434N2420  
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

## FEATURES

source  
 1..2297  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="RZPD:DKFp434N2420"  
 /db\_xref="taxon:9606"  
 /clone="DKFp434N2420"  
 /issue\_type="testis"  
 /clone\_id="434 (synonym: hte03). Vector pSport1, host DH10B; sites NotI + SalI"  
 /dev\_stage="adult"  
 /note="cell cycle checkpoint protein CHFR, N-terminus

truncated"  
 1..2297  
 /gene="DKFp434N2420"  
 <1..921  
 /gene="DKFp434N2420"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA870812.1"  
 /db\_xref="GI:6808266"  
 /translation="VEAAILIQHPDKSRSEEDVOSMDARNKITQDMLQPKVRSFSDSE  
 GSEDLLELDVDSSESDISQPYVVCQCEYERAOAPHCAPBEPBEPAPALGPA  
 PSTRSVLITVAVDVPVCPLOGSHALCTCFQPMRPRARERODPRVAPQCAVCQPC  
 HLYMGCTRTCTGCTAFCGLNIGDKLIDVANNNSYSDIILKNYLATRGLTWKMLT  
 ESLVALORGVFLSDIRVTDYTLCTCCGRSRELTYQIRQNI PASELPVATYSRD  
 CYWGRNCRITYVKAHAKFNHICEITFKN"

## ORIGIN

Query Match 55.9%; Score 1497; DB 3; Length 2297;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1165	GTGGAAGCATACCTCATTCAGACATCCAGACAGAGATGCGAGTGAAGAAGATGCAAAAGT	1224
DB	1	GTGGAAGCATACCTCATTCAGACATCCAGACAGAGATGCGAGTGAAGAAGATGCAAAAGT	60
QY	1225	ATGATGCGCAGGAATTAATCACTCAAGACATGCTGCGACCCCAAGTCAAGCGGCTCTTT	1284
DB	61	ATGATGCGCAGGAATTAATCACTCAAGACATGCTGCGACCCCAAGTCAAGCGGCTCTTT	120
QY	1285	TCTGATGAAGAAAGGAGTTGAGAGACCTGCTGAGCTGTGACAGTTGACAGTAGTCC	1344
DB	121	TCTGATGAAGAAAGGAGTTGAGAGACCTGCTGAGCTGTGACAGTTGACAGTAGTCC	180
QY	1345	TCAGACATTAGCCAGCCATACGTGCTGCGGAGATGCTCTGAGTACAGAAAGCGAGCG	1404
DB	181	TCAGACATTAGCCAGCCATACGTGCTGCGGAGATGCTCTGAGTACAGAAAGCGAGCG	240
QY	1405	GCGCAGCTCTCCCACTGCGCCAGCACCCGAGGGCCAGAGAGCCCAAGGCCCTGGGG	1464
DB	241	GCGCAGCTCTCCCACTGCGCCAGCACCCGAGGGCCAGAGAGCCCAAGGCCCTGGGG	300
QY	1465	GATGACACCTCCAGCGTCCGTGAGCTGAGCAGACAGAGTTCAGAGTTACGTGTGCTCTG	1524
DB	301	GATGACACCTCCAGCGTCCGTGAGCTGAGCAGACAGAGTTCAGAGTTACGTGTGCTCTG	360
QY	1525	CAAGGAAGCCAGCGCTGTGCACTGCTGCTTCCAGCCATGCCCCGAGACCGAGAGCGAG	1584
DB	361	CAAGGAAGCCAGCGCTGTGCACTGCTGCTTCCAGCCATGCCCCGAGACCGAGAGCGAG	420
QY	1585	CGCGAGCAGACCCGCGTGTGCGCCCTCAAGAGTGTGGGTCTGCTGACGCTTCTGCG	1644
DB	421	CGCGAGCAGACCCGCGTGTGCGCCCTCAAGAGTGTGGGTCTGCTGACGCTTCTGCG	480
QY	1645	CACCTGTACTGGGGGTGACACCGGACCGGCTGCTAAGCGCTGCGCCCGTTTGTGAG	1704
DB	481	CACCTGTACTGGGGGTGACACCGGACCGGCTGCTAAGCGCTGCGCCCGTTTGTGAG	540
QY	1705	CTCAACTGGGTGACAGATGCTGTGACCGCGTGTCTGAACAACAAGCTACAGATCAAC	1764
DB	541	CTCAACTGGGTGACAGATGCTGTGACCGCGTGTCTGAACAACAAGCTACAGATCAAC	600
QY	1765	ATCCGTAAGANTTACTCTGGCAACGAGAGTTTGACATGGAAGAAACATGTTGACGAGAGC	1824
DB	601	ATCCGTAAGANTTACTCTGGCAACGAGAGTTTGACATGGAAGAAACATGTTGACGAGAGC	660
QY	1825	CTCGTGGCTCTCAACGGGGAGTGTCTGCTGTCTGATTAACAAGTACAGGAGACACC	1884
DB	661	CTCGTGGCTCTCAACGGGGAGTGTCTGCTGTCTGATTAACAAGTACAGGAGACACC	720
QY	1885	GTTCTGTGTTACTGCTGTGAGCTGTGAGCTTCCGTGAGCTGACCTATCAAGTATGGGAG	1944
DB	721	GTTCTGTGTTACTGCTGTGAGCTGTGAGCTTCCGTGAGCTGACCTATCAAGTATGGGAG	780



QY 1945 AACATTCTGCTTCGAGTTGCGAGTGGCCGTTAATCCCTCTGACTGCTACTG393G 2004  
 DB 781 AACATTCTGCTTCGAGTTGCGAGTGGCCGTTAATCCCTCTGACTGCTACTG393G 840  
 QY 2005 CGTAACTGCGGCACTGAGTGAAGTCAACGCGCATGAAATTCATATCTGTGA 2064  
 DB 841 CGTAACTGCGGCACTGAGTGAAGTCAACGCGCATGAAATTCATATCTGTGA 900  
 QY 2065 CAGACAGGTTCAAAAATTAAGCATCAAGAGCCCTGAGAGCTTCAAGCATGAGT 2124  
 DB 901 CAGACAGGTTCAAAAATTAAGCATCAAGAGCCCTGAGAGCTTCAAGCATGAGT 960  
 QY 2125 AAGAGAGCGTGTGTTTAAATACAGACAGACAGCTCAAGTGTTCACAGCCCCCTG 2184  
 DB 961 AAGAGAGCGTGTGTTTAAATACAGACAGACAGCTCAAGTGTTCACAGCCCCCTG 1020  
 QY 2185 AAGGAAAGGACGAGAGGCTCTCCGACAGTGTCTTGGGGTGAATCTTCTGTGAGCTTTT 2244  
 DB 1021 AAGGAAAGGACGAGAGGCTCTCCGACAGTGTCTTGGGGTGAATCTTCTGTGAGCTTTT 1080  
 QY 2245 ACCCTGTGAGTGAACCTCCCGAGAGCCCGGAGCCCGGCTCTGTGAGG 2304  
 DB 1081 ACCCTGTGAGTGAACCTCCCGAGAGCCCGGAGCCCGGCTCTGTGAGG 1140  
 QY 2305 GCTGGGCAAGGCTCTGTGAGTGAACAGACAGACAGCTTCTGTGAATGCGG 2364  
 DB 1141 GCTGGGCAAGGCTCTGTGAGTGAACAGACAGACAGCTTCTGTGAATGCGG 1200  
 QY 2365 CGTCCGCGGAGAGGGGCAAGTGTCTTGTGATCAATTTCCGAACTCAAGTTAAGC 2424  
 DB 1201 CGTCCGCGGAGAGGGGCAAGTGTCTTGTGATCAATTTCCGAACTCAAGTTAAGC 1260  
 QY 2425 AGAAGCTGTGTTTCAAGAAAGTTTCAAGGAGAGAGGCAAGTTTATCAAAATCTGT 2484  
 DB 1261 AGAAGCTGTGTTTCAAGAAAGTTTCAAGGAGAGAGGCAAGTTTATCAAAATCTGT 1320  
 QY 2485 TCAGAGAGAGGAGCATTAAGTTTACAGCTCAAGACAGTCAATATCTGTGCTGG 2544  
 DB 1321 TCAGAGAGAGGAGCATTAAGTTTACAGCTCAAGACAGTCAATATCTGTGCTGG 1380  
 QY 2545 AAAACCAAGCATTTTATCTATTTTATTAATAGTTTGTGCTTATCTTAAATA 2604  
 DB 1381 AAAACCAAGCATTTTATCTATTTTATTAATAGTTTGTGCTTATCTTAAATA 1440  
 QY 2605 GATTTAAATGTCAAACTGTAGCAAAATTAATTAATTAATTAATTAATTAATGAC 2661  
 DB 1441 GATTTAAATGTCAAACTGTAGCAAAATTAATTAATTAATTAATTAATTAATGAC 1497

RESULT 2  
 AL539798 990 bp mRNA linear EST 24-MAR-2004  
 LOCUS AL539798 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
 DEFINITION CS0DF031YA23 5-PRIME, mRNA sequence.  
 ACCESSION AL539798  
 VERSION AL539798.3 GI:45715436  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 990)  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:31264361.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequenage  
 BP 191 91006 Evry cedex - France  
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo (dr) primer. Five prime  
 end enriched, double-strand cDNA was digested with NotI and cloned  
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library

was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6792.r  
 For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?c=CS0DF031YA23&pic=6792.r>.  
 Location/Qualifiers  
 1..990  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF031YA23"  
 /issue\_type="FETAL BRAIN"  
 /dev\_stage="fetal"  
 /clone\_1ib="Homo sapiens FETAL BRAIN"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo (dr) primer. Five prime end  
 enriched, double-strand cDNA was digested with NotI and  
 cloned into the NotI and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Query Match 30.8%; Score 825; DB 1; Length 990;  
 Best Local Similarity 100.0%; Pred.No. 0;  
 Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GTTCGGCGGGGCGGAGTGAATCCCGATGAGCGGCGGAGGAGGCAAGCAGTGC 120  
 DB 20 GTTCGGCGGGGCGGAGTGAATCCCGATGAGCGGCGGAGGAGGCAAGCAGTGC 79  
 QY 121 CGCGCGCGGAGCGGCTCTGCGTCTGGGGCGGAGGAGGCGAGCCGCGAC 180  
 DB 80 CGCGCGCGGAGCGGCTCTGCGTCTGGGGCGGAGGAGGCGAGCCGCGAC 139  
 QY 181 GTTCCTCTGAGGAGAGCGGAGTGAACATCGGCGGAGAGAGTTCGACCTTCTTC 240  
 DB 140 GTTCCTCTGAGGAGAGCGGAGTGAACATCGGCGGAGAGAGTTCGACCTTCTTC 199  
 QY 241 CCAGCAATTAATCTGCTCTGAGATCACTGAATGATGATGATGATGATGATGATGAT 300  
 DB 200 CCAGCAATTAATCTGCTCTGAGATCACTGAATGATGATGATGATGATGATGATGAT 259  
 QY 301 CAGGTGACACTGGAAGATACAGCAACAGTGAACAGTGAATTAACAGTGAAGTTGT 360  
 DB 260 CAGGTGACACTGGAAGATACAGCAACAGTGAACAGTGAATTAACAGTGAAGTTGT 319  
 QY 361 AAGAGCAGACATGCCCTTTACAGACTGGGAGTGCATCTTGTGTTACAGAAAT 420  
 DB 320 AAGAGCAGACATGCCCTTTACAGACTGGGAGTGCATCTTGTGTTACAGAAAT 379  
 QY 421 GAACGGGAGCAACAGTGCATACCTCTTAATCTTTAATGTAAGAAAGCAAGGCAATGAC 480  
 DB 380 GAACGGGAGCAACAGTGCATACCTCTTAATCTTTAATGTAAGAAAGCAAGGCAATGAC 439  
 QY 481 CAGGAATCTTTGAAGTCAAGGAAATGTTTCATGAGGACCAAAATATCTCAGGT 540  
 DB 440 CAGGAATCTTTGAAGTCAAGGAAATGTTTCATGAGGACCAAAATATCTCAGGT 499  
 QY 541 GAGGTGACGGGCGAGGGCCGATCCCGGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 DB 500 GAGGTGACGGGCGAGGGCCGATCCCGGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 559  
 QY 601 TGGTTGAGGAGCAACAGCACAATCAATGAGAGTCAAGCTTTCCCAACAGCTCGGCG 660  
 DB 560 TGGTTGAGGAGCAACAGCACAATCAATGAGAGTCAAGCTTTCCCAACAGCTCGGCG 619  
 QY 661 TCTTCAAGGAGCTTCTCTGAGGCGAGAGCTTCTCAAGTTGTTGAGTCTGAGGAGT 720  
 DB 620 TCTTCAAGGAGCTTCTCTGAGGCGAGAGCTTCTCAAGTTGTTGAGTCTGAGGAGT 679  
 QY 721 GGTGATCTCCCTTAAGAAAGTGTCTCTGTGGCAAGTGAATGAATCTTCAAGCTTT 780  
 DB 680 GGTGATCTCCCTTAAGAAAGTGTCTCTGTGGCAAGTGAATGAATCTTCAAGCTTT 739

QY 781 GCTCAGCTCTCCAGACAGAAAGACTGCTCTTTCCTGTTGAACCCAGATCAG 840  
 DB 740 GCTCAGCTCTCCAGACAGAAAGACTGCTCTTTCCTGTTGAACCCAGATCAG 799  
 QY 841 GAGATTTGAGCCCGTGAAGAAATAAGAGAGATGAGGAC 885  
 DB 800 GAGATTTGAGCCCGTGAAGAAATAAGAGAGATGAGGAC 844

RESULT 3  
 LOCUS BU176559 921 bp mRNA linear EST 04-SEP-2002  
 DEFINITION AGENCOURT\_7940317 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:6143609  
 5', mRNA sequence.  
 ACCESSION BU176559  
 VERSION BU176559.1 GI:22690543  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 921)  
 NIH-MGC http://mgi.mgi.nhl.gov/.  
 Mammalian Gene Collection (MGC)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nhl.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
 http://image.llnl.gov  
 Plate: L1AM13467 row: e column: 18  
 High quality sequence stop: 697.

FEATURES  
 source location/Qualifiers  
 1..921  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6143609"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pCMV-Sport6; Site: 1: NotI;  
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."

ORIGIN  
 Query Match 27.0%; Score 722; DB 5; Length 921;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1889 TGTGTACTGCTGTGGCTGCGCAGTTTCCGAGTCACTCATGTTGGCAGACA 1948  
 DB 1 TGTGTACTGCTGTGGCTGCGCAGTTTCCGAGTCACTCATGTTGGCAGACA 60  
 QY 1949 TTCCTGCTCCGAGTTGGCAGTGGCGTGAATCCCGTCTGACTGCTCTGAGGCGCGTA 2008  
 DB 61 TTCCTGCTCCGAGTTGGCAGTGGCGTGAATCCCGTCTGACTGCTCTGAGGCGCGTA 120  
 QY 2009 ACTGCGCAGCTCAGGTGAAGTCAACACGCCATGTAATCATATCTGTGAACAGA 2068  
 DB 121 ACTGCGCAGCTCAGGTGAAGTCAACACGCCATGTAATCATATCTGTGAACAGA 180  
 QY 2069 CAAAGTTCAAAATCTAAGCATCCAGAGCCCTGAGCAGCTTTCAGCACTGAGGTGAAGA 2128  
 DB 181 CAAAGTTCAAAATCTAAGCATCCAGAGCCCTGAGCAGCTTTCAGCACTGAGGTGAAGA 240  
 QY 2129 GAGCGGTGTTTAAATATAGAGACAAGACGTCAGGTGTTTTCAGAGCCCTGAGGAG 2188

DB 241 GAGCGTGTGTTTAAATATAGAGACAAGACAGTCAAGGTGTTTTCAGAGCCCTGAGG 300  
 QY 2189 AAGGAGCCAGAGGTCTCCGACAGGTGCTGTGGGTGATCTTTCTGTGAGCTTTTACC 2248  
 DB 301 AAGGAGCCAGAGGTCTCCGACAGGTGCTGTGGGTGATCTTTCTGTGAGCTTTTACC 360  
 QY 2249 TCTGATGAGACCCCTCCAGAGCCCGGGGGCGCAGCCCGCTCTGTGAGCGCTG 2308  
 DB 361 TCTGATGAGACCCCTCCAGAGCCCGGGGGCGCAGCCCGCTCTGTGAGCGCTG 420  
 QY 2309 GGCAGAGGCTCGTGTGAGTCAAGCAGACAGCAAGCTTCTGTGAATATGAGCGCTG 2368  
 DB 421 GGCAGAGGCTCGTGTGAGTCAAGCAGACAGCAAGCTTCTGTGAATATGAGCGCTG 480  
 QY 2369 CCGCCGAGAGGGCGAGTTTGTCTTTTGTATCTTTCCGAAACTGACGTTAAAGACAGA 2428  
 DB 481 CCGCCGAGAGGGCGAGTTTGTCTTTTGTATCTTTCCGAAACTGACGTTAAAGACAGA 540  
 QY 2429 GTCGTGTTTCAAGAAAGTTCAAGGGAGAAAGGCAAGTTATCAAAACATGTTGCG 2488  
 DB 541 GTCGTGTTTCAAGAAAGTTCAAGGGAGAAAGGCAAGTTATCAAAACATGTTGCG 600  
 QY 2489 GAGAAGGAGCATAGTTTACAGCTACAGACGATACACAATATCTGCTGCGGAAA 2548  
 DB 601 GAGAAGGAGCATAGTTTACAGCTACAGACGATACACAATATCTGCTGCGGAAA 660  
 QY 2549 CCACAGCATTTATCTATTTTATTTAATAGTTGCTGCTATCTTCTAATAGATT 2608  
 DB 661 CCACAGCATTTATCTATTTTATTTAATAGTTGCTGCTATCTTCTAATAGATT 720  
 QY 2609 TAAATGTACAAACGTGACCAAAATATATTAATTTTCAAAATGAC 2661  
 DB 721 TAAATGTACAAACGTGACCAAAATATATTAATTTTCAAAATGAC 773

RESULT 4  
 LOCUS B0073354 992 bp mRNA linear EST 02-APR-2002  
 DEFINITION AGENCOURT\_7052199 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:5806143  
 5', mRNA sequence.  
 ACCESSION B0073354  
 VERSION B0073354.1 GI:19902400  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 992)  
 NIH-MGC http://mgi.mgi.nhl.gov/.  
 Mammalian Gene Collection (MGC)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nhl.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at:  
 http://image.llnl.gov  
 Plate: L1CM2046 row: h column: 16  
 High quality sequence stop: 670.

FEATURES  
 source location/Qualifiers  
 1..992  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5806143"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone="NIH MGC 101"  
 /note="Organ: lung; Vector: pORG7; Site: 1: EcoRI; Site: 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned

**ORIGIN**

Matches 70

Db 661 AAGATGTGCAAGTATGGATGCCAGGAATAAATCACTCAAGA 703

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo..

**COMMENT**

High quality sequence stop: 660.

**Source**

11

Query Match

**Matches** 74

1527 QY

Dh

TEST

2

1

1007

DB 184

QY 1767

Db 244

1827  
QY

**Db 304**

1887 ov

354

?

...

204

2067  
57

Db 544 GACAAAGTTCAAAAATAAGCATCCAGAGAGCCCTGAGCAGCTTTTCAAGCATGAGAGTGA 603

QY 2127 GAGAGCGTGTCTTTTAAATATACAGAGCAAGCATGAGCTTTTTCACAGCCCTCTGAG 2186

Db 604 GAGAGCGTGTCTTTTAAATATACAGAGCAAGCATGAGCTTTTTCACAGCCCTCTGAG 663

QY 2187 GGAAGGAGCAGAGGCTCTCCGACAGAGTCTTGGGGGTGACCTTTCTGTGAGAGCTTTTAC 2246

Db 664 GGAAGGAGCAGAGGCTCTCCGACAGAGTCTTGGGGGTGACCTTTCTGTGAGAGCTTTTAC 723

QY 2247 CCTCTGAGTGAAGCCTCCCGAGAGGCC 2275

Db 724 CCTCTGAGTGAAGCCTCCCGAGAGGCC 752

RESULT 5  
LOCUS CR625905 2412 bp mRNA linear HTC 21-JUL-2004  
DEFINITION Full-length cDNA clone CS0DF01Y119 of Fetal brain of Homo sapiens (human).  
ACCESSION CR625905  
VERSION CR625905.1 GI:50506712  
KEYWORDS HTC, CNSLT cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 2412)  
AUTHORS Li W.B., Gruber C., Jessee J., and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
CONTACT: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue  
2 (bases 1 to 2412)  
REFERENCE 2 (bases 1 to 2412)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
FEATURES  
source location/Qualifiers  
1..2412  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF01Y119"  
/issue\_type="Fetal brain"  
/plasmid="pCMVSPORT\_6"

ORIGIN  
Query Match 24.5%; Score 657; DB 3; Length 2412;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 887; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1772 AGAATTACCTGACAGACAGAGTTTGAATGAAAAACATGTTGACCGAGAGGCTGTGG 1831

Db 950 AGAATTACCTGACAGACAGAGTTTGAATGAAAAACATGTTGACCGAGAGGCTGTGG 1009

QY 1832 CTCCTCAGCGGGAGAGTCTCTGCTGTGATTACAGAGTCAAGGAGACACCGTTCTGT 1891

Db 1010 CTCCTCAGCGGGAGAGTCTCTGCTGTGATTACAGAGTCAAGGAGACACCGTTCTGT 1069

QY 1892 GTTACTGCTGTGGCTTGGCGGAGCTTCCGTGACCTATGATGATCGGACAGAAATTC 1951

Db 1070 GTTACTGCTGTGGCTTGGCGGAGCTTCCGTGACCTATGATGATCGGACAGAAATTC 1129

QY 1952 CTGCTTCCGAGTTGCAAGTGGCGGTAAATCCGCTCTGCTCTGAGGCGGTAACT 2011

Db 1130 CTGCTCCGAGTTGCGCAGTGGCGGTAAATCCCGTCTGACTGCTACTGAGGCGGTAACT 1189

QY 2012 GCCGCAGTCAAGTGAAGAGTCCACAGCCGATGAATTCATATCTGTGAACAGACAA 2071

Db 1190 GCCGCAGTCAAGTGAAGAGTCCACAGCCGATGAATTCATATCTGTGAACAGACAA 1249

QY 2072 GTTCAAAAATAAGCATCCAGAGGCGCTTGAAGAGCTTTTCAAGAGTGAAGAGAG 2131

Db 1250 GTTCAAAAATAAGCATCCAGAGGCGCTTGAAGAGCTTTTCAAGAGTGAAGAGAG 1309

QY 2132 CGTGTCTTAAATATACAGAGCAAGCATGTCAGAGTGTCTTCAAGAGTGAAGAG 2191

Db 1310 CGTGTCTTAAATATACAGAGCAAGCATGTCAGAGTGTCTTCAAGAGTGAAGAG 1369

QY 2192 GAGCGAGGCTCTCCGACAGTCTCTGGGGTGAATCTTCTGTGAGAGCTTTTACCTCT 2251

Db 1370 GAGCGAGGCTCTCCGACAGTCTCTGGGGTGAATCTTCTGTGAGAG--TTTACCTCT 1427

QY 2252 GAGTGAAGCCTCCCGAGAGCCCGGGGGCGGAGCCCGCTCTGAGAGCGCTGGAG 2311

Db 1428 GAGTGAAGCCTCCCGAGAGCCCGGGGGCGGAGCCCGCTCTGAGAGCGCTGGAG 1487

QY 2312 AGGCTCTGTGTGATCAGAGAGAGAGCAGAGCAGTCTTGTATACATGCGGCGCTCCG 2371

Db 1488 AGGCTCTGTGTGATCAGAGAGAGAGCAGAGCAGTCTTGTATACATGCGGCGCTCCG 1547

QY 2372 CCGAGAGGGGAGATTTTCTCTTTTGTATCATTTTCCGAACTACAGTTAAAGAGAGTC 2431

Db 1548 CCGAGAGGGGAGATTTTCTCTTTTGTATCATTTTCCGAACTACAGTTAAAGAGAGTC 1607

QY 2432 TGTCTTCAAGAAAGTTTCAAGGGAGAGGCAAGTTATCAAAAACATGTTTCAGAG 2491

Db 1608 TGTCTTCAAGAAAGTTTCAAGGGAGAGGCAAGTTATCAAAAACATGTTTCAGAG 1667

QY 2492 AAGGAGCATATAGTTTACAGCTACAGAGCAGTACCAATATCTGCTGAGGAAACCA 2551

Db 1668 AAGGAGCATATAGTTTACAGCTACAGAGCAGTACCAATATCTGCTGAGGAAACCA 1727

QY 2552 CAGCATTTTATCTATTTTATTTTAAAGTTTGTGCTTATCTTCTTAAATGAATTTAA 2611

Db 1728 CAGCATTTTATCTATTTTATTTTAAAGTTTGTGCTTATCTTCTTAAATGAATTTAA 1787

QY 2612 ATGTCAAAAGTGTGACCAATATATATATTTATTTTCAAAATGAC 2661

Db 1788 ATGTCAAAAGTGTGACCAATATATATTTATTTTCAAAATGAC 1837

RESULT 7  
BF971929 897 bp mRNA linear EST 22-JAN-2001  
LOCUS BF971929  
DEFINITION M62240332F1 NIH\_MGC\_46 Homo sapiens cDNA clone IMAGE:4329079 5', mRNA sequence.  
ACCESSION BF971929  
VERSION BF971929.1 GI:12339144  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 897)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC), Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNI at: http://image.llnl.gov  
Plate: LCM1189 row: p column: 08

## FEATURES High quality sequence stop: 728.

Location/Qualifiers

1.897

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4329079"

/issue\_type="leiomyosarcoma cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_46"

/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into pCR2.1/KS+ vector using the following 5' adaptor: GGAGGAG(G). Size-selected &gt;500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 24.2%; Score 647; DB 2; Length 897;  
Best Local Similarity 100.0%; Pred. No. 1.6e-311;  
Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1908 GGGAGCTCCGAGCTGACCTATGATATGGCAGAAACATTCCTCCGAGTTGCC 1967  
1 GGGAGCTCCGAGCTGACCTATGATATGGCAGAAACATTCCTCCGAGTTGCC 60  
1968 AGTGGCGGTAAACATCCGCTCTGACTGCTACTGCGGCGGTAACTGCGGCACTCAGGTGAA 2027  
61 AGTGGCGGTAAACATCCGCTCTGACTGCTACTGCGGCGGTAACTGCGGCACTCAGGTGAA 120  
2028 AGTCCACACGCGGTAAATTCATATCTGTGAACAGACAGGTTCAAAATCTAACG 2087  
121 AGTCCACACGCGGTAAATTCATATCTGTGAACAGACAGGTTCAAAATCTAACG 180  
2088 ATCCAGAGGCGCGGAGAGCTTCAACACCTGAGAGGAGAGGCGCTTTTAAATAC 2147  
181 ATCCAGAGGCGCGGAGAGCTTCAACACCTGAGAGGAGAGGCGCTTTTAAATAC 240  
2148 AGAGACAGACAGCTCAAGGTGTTTCAAGCCCCCTGAGGAGAGGAGCCAGGCTCCG 2207  
241 AGAGACAGACAGCTCAAGGTGTTTCAAGCCCCCTGAGGAGAGGAGCCAGGCTCCG 300  
2208 AAGGTGCTCTGAGGAGTCTTCTGTGAGCTTTTACCTCTGAGTGAACCTCTCCC 2267  
301 AAGGTGCTCTGAGGAGTCTTCTGTGAGCTTTTACCTCTGAGTGAACCTCTCCC 360  
2268 AAGAGCGCGGAGGCGGAGCGGCGCTCTGAGTGAACGCTGAGGAGGCTCTGTGTGGA 2327  
361 AAGAGCGCGGAGGCGGAGCGGCGCTCTGAGTGAACGCTGAGGAGGCTCTGTGTGGA 420  
2328 TCAGCAGCAGACGAGAGCTTCTGTAAACATGCGGCGTCCGCGAGAGGGGAGTT 2387  
421 TCAGCAGCAGACGAGAGCTTCTGTAAACATGCGGCGTCCGCGAGAGGGGAGTT 480  
2388 TGTCTTTTGTACATTTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTTCAGAGAAAGT 2447  
481 TGTCTTTTGTACATTTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTTCAGAGAAAGT 540  
2448 TTCAGAGGAGAGAGGAGGAGTTTATCAAAACATGTTTTCAGAGAGAGGAGTAAAGTT 2507  
541 TTCAGAGGAGAGAGGAGGAGTTTATCAAAACATGTTTTCAGAGAGAGGAGTAAAGTT 600  
2508 ACAGCTTACAGAGCTACCAATATCTGCTGTGAGAGAGAGAGAGAG 2554  
601 ACAGCTTACAGAGCTACCAATATCTGCTGTGAGAGAGAGAGAGAG 647

RESULT 8  
CN428709  
LOCUS CN428709 755 bp mRNA linear EST 16-MAY-2004

## DEFINITION 17000600090657 GRN\_PRENED Homo sapiens cDNA 5', mRNA sequence.

Accession CN428709  
Version CN428709.1 GI:47416303

KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

Homo sapiens

ORIGIN

Query Match 23.9%; Score 640; DB 7; Length 755;  
Best Local Similarity 100.0%; Pred. No. 5.2e-308;  
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2022 GGTGAAGCTCACACGACCATTAATATATCTGTGAACAGACAGAGTTCAAAA 2081  
1 GGTGAAGCTCACACGACCATTAATATATCTGTGAACAGACAGAGTTCAAAA 60  
2082 CTAAAGATTCAGAGGCGCTGAGCAGCTTCACTGAGAGTGAAGAGCGTGTGTTTA 2141  
61 CTAAAGATTCAGAGGCGCTGAGCAGCTTCACTGAGAGTGAAGAGCGTGTGTTTA 120  
2142 AATATCAGAGACAGACAGCTCAAGGTGTTTCAAGCCCCCTGAGGAGAGGAGCGCAGG 2201  
121 AATATCAGAGACAGACAGCTCAAGGTGTTTCAAGCCCCCTGAGGAGAGGAGCGCAGG 180  
2202 TCTCCGACAGGTGCTCTGAGGAGTCTTCTGTGAGCTTTTACCTCTGAGTGAAC 2261  
181 TCTCCGACAGGTGCTCTGAGGAGTCTTCTGTGAGCTTTTACCTCTGAGTGAAC 240  
2262 CTCCGAGAGCGCGGAGGCGGAGCGGCGCTCTGAGTGAACGCTGAGGAGGCTCTG 2321  
241 CTCCGAGAGCGCGGAGGCGGAGCGGCGCTCTGAGTGAACGCTGAGGAGGCTCTG 300  
2322 GTGGCATCAGCAGCAGACGAGAGCTTCTGTAAACATGCGGCGTCCGCGAGAGGGG 2381  
301 GTGGCATCAGCAGCAGACGAGAGCTTCTGTAAACATGCGGCGTCCGCGAGAGGGG 360  
2382 CAGTTTGTCTTTTGTACATTTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTCAGG 2441  
361 CAGTTTGTCTTTTGTACATTTTCCGAAACTACAGTTAAAGCAGAGTCTGTTTCAGG 420  
2442 AAAAGTTTCAAGGAGAGAGGAGGAGTTTATCAAAACATGTTTTCAGAGAGAGGAGAT 2501  
421 AAAAGTTTCAAGGAGAGAGGAGGAGTTTATCAAAACATGTTTTCAGAGAGAGGAGAT 480





ORIGIN

/cissue\_type="placenta"  
/clone\_lib="PLACB1"  
/note=Vector: PME18SFL3"

Query Match 22.4%; Score 600; DB 1; Length 680;  
Best Local Similarity 100.0%; Pred. No. 5.2e-288;  
Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1503 CCAAGATTACGTGCTCCCTCTGCAAGAAAGCAAGCCCTGCTGCACTGCTGCTTCCAGCC 1562  
DB 1 CCAAGATTACGTGCTCCCTCTGCAAGAAAGCAAGCCCTGCTGCACTGCTGCTTCCAGCC 60  
QY 1563 CATTGCCGAGCCGAGAGCGAGAGCGAGAGAGCCGCGTGTGCGCCCTCAGAGAGTGC 1622  
DB 61 CATTGCCGAGCCGAGAGCGAGAGCGAGAGAGCCGCGTGTGCGCCCTCAGAGAGTGC 120  
QY 1623 GGTCTGCTGCAAGCTTCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 1682  
DB 121 GGTCTGCTGCAAGCTTCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 180  
QY 1683 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1742  
DB 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 1743 CAACACAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 1802  
DB 241 CAACACAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 300  
QY 1803 GAAAAACATGTTGACCGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1862  
DB 301 GAAAAACATGTTGACCGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 1863 TTACAGAGTCAAGCGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1922  
DB 361 TTACAGAGTCAAGCGAGAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 1923 GGTGACCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 1982  
DB 421 GGTGACCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 480  
QY 1983 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2042  
DB 481 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 2043 GAAATTCATCATATCTGTAAGACAGCAAGGTTCAAAAATAAGATTCAGAGCCCTGA 2102  
DB 541 GAAATTCATCATATCTGTAAGACAGCAAGGTTCAAAAATAAGATTCAGAGCCCTGA 600

RESULT 11  
BQ215731 870 bp mRNA linear EST 02-MAY-2002  
LOCUS AGENCOURT 7549119 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6059147  
DEFINITION 5', mRNA Sequence.  
ACCESSION BQ215731  
VERSION BQ215731.1 GI:20397131  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 870)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaab@remail.nih.gov  
Tissue Procurement: DCTD/DIR/Gardar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LNL13325 row: f column: 12  
High quality sequence stop: 601.  
Location/Qualifiers  
1..870  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6059147"  
/cissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 68"  
/note="Organ: lung; Vector: pCW-SORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.8 Kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 21.8%; Score 583; DB 5; Length 870;  
Best Local Similarity 100.0%; Pred. No. 1.6e-279;  
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 CAACCGCTGAAATGCGCAACCGTCCAGAGAGCTGAGAGAGCGGCTGGAAAGCA 975  
DB 1 CAACCGCTGAAATGCGCAACCGTCCAGAGAGCTGAGAGAGCGGCTGGAAAGCA 60  
QY 976 GACAAAGTGAAGAGACGCTGACATGATCATCTGCAAGACCTGCTGCAAGCTGCTG 1035  
DB 61 GACAAAGTGAAGAGACGCTGACATGATCATCTGCAAGACCTGCTGCAAGCTGCTG 120  
QY 1036 AGTTTGACGCTGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1095  
DB 121 AGTTTGACGCTGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 1096 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1155  
DB 181 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 1156 AACAACTGCTGAAAGATCACTCATTCAGATTCAGACAGAGTTCAGTGAAGAT 1215  
DB 241 AACAACTGCTGAAAGATCACTCATTCAGATTCAGACAGAGTTCAGTGAAGAT 300  
QY 1216 GTGCAAGTGAAGTGCAGAAATTAATCATCTCAAGACATCTGCAAGCCCAAGTCAAG 1275  
DB 301 GTGCAAGTGAAGTGCAGAAATTAATCATCTCAAGACATCTGCAAGCCCAAGTCAAG 360  
QY 1276 CGGTCTTTTCTGATGAAGAGAGGATTGAGAGACTGCTGAGCTGCAAGCTTGAC 1335  
DB 361 CGGTCTTTTCTGATGAAGAGAGGATTGAGAGACTGCTGAGCTGCAAGCTTGAC 420  
QY 1336 AGTAGTCTCAGACATTAGCAGGCAATACGTGCTGCTGCTGCTGCTGCTGCTGCTG 1395  
DB 421 AGTAGTCTCAGACATTAGCAGGCAATACGTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 1396 AGGCAAGCGGCGCAGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1455  
DB 481 AGGCAAGCGGCGCAGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 1456 GCCCTGGGGAGTCAACCTTCAGAGTCCGCTGAGCTGAGAGAG 1498  
DB 541 GCCCTGGGGAGTCAACCTTCAGAGTCCGCTGAGAGAG 583

RESULT 12  
BM459647 916 bp mRNA linear EST 05-FEB-2002  
LOCUS AGENCOURT 6418058 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5534277  
DEFINITION 5', mRNA Sequence.  
ACCESSION BM459647  
VERSION BM459647.1 GI:18508687  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/  
 1 (bases 1 to 916)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rcmail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: LLM12320 row: h column: 22  
 High quality sequence stop: 642.  
 Location/Qualifiers

FEATURES  
 source  
 1..916  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5534277"  
 /issue\_type="leiomysarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 71"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; NotI;  
 site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.1 kb."

## ORIGIN

Query Match 21.1%; Score 565; DB 4; Length 916;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-270;  
 Matches 565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 709 GGGTCTGGGGGCTGGTGGCATCTCCCTAAAGAGTGGTCTCTGTGGCAAGTATGA 768  
 155 GGGTCTGGGGGCTGGTGGCATCTCCCTAAAGAGTGGTCTCTGTGGCAAGTATGA 214  
 769 GTCCTCAGCTTGGCTCAGCTCTCCAGACAGAAAGATGCGCTCTTTCGTCTGGAA 828  
 215 GTCCTCAGCTTGGCTCAGCTCTCCAGACAGAAAGATGCGCTCTTTCGTCTGGAA 274  
 829 CCCAGGATCAGAGAGATTTGGAGCCCGTGAAGAGAAATGAGAGATGGGACCTT 888  
 275 CCCAGGATCAGAGAGATTTGGAGCCCGTGAAGAGAAATGAGAGATGGGACCTT 334  
 889 GACCTGAACGGGAGTGTGTGGTGCACAAACCGCTGAAGAAATGCCAAACCGTCCACG 948  
 335 GACCTGAACGGGAGTGTGTGGTGCACAAACCGCTGAAGAAATGCCAAACCGTCCACG 394  
 949 GACCTGAACGGGAGTGTGTGGTGCACAAACCGCTGAAGAAATGCCAAACCGTCCACG 1008  
 395 GACCTGAACGGGAGTGTGTGGTGCACAAACCGCTGAAGAAATGCCAAACCGTCCACG 454  
 455 TGCAGAGACCTGCTGACAGCTGCTGAGTTTCAACCCCTGACATGACAGTTCTGCGG 514  
 1009 TGCAGAGACCTGCTGACAGCTGCTGAGTTTCAACCCCTGACATGACAGTTCTGCGG 1068  
 515 GCTTGTCTACTCGGGCTGGATGAGCGCTGCTGCTGTGCTTACCTGCGCTGTCCCGT 574  
 1069 GCTTGTCTACTCGGGCTGGATGAGCGCTGCTGCTGTGCTTACCTGCGCTGTCCCGT 1128  
 515 GCTTGTCTACTCGGGCTGGATGAGCGCTGCTGCTGTGCTTACCTGCGCTGTCCCGT 574  
 1129 GAGCGGATCTGTAAAAACCATCTCAACAACTGTGGAGATACCTCACTCCAGCAT 1188  
 575 GAGCGGATCTGTAAAAACCATCTCAACAACTGTGGAGATACCTCACTCCAGCAT 634  
 1189 CCAGACAAGTGTGCAAGTGAAGAGATGTGCAAGATGTGATGCCAGATTAATCACT 1248  
 635 CCAGACAAGTGTGCAAGTGAAGAGATGTGCAAGATGTGATGCCAGATTAATCACT 694  
 1249 CAAGACATGCTGACGCCCAAGTCA 1273

Db 695 CAAGACATGCTGACGCCCAAGTCA 719

RESULT 13  
 BMS42545 1143 bp mRNA linear EST 20-FEB-2002  
 LOCUS AGENCOURT 6426028 NIH\_MGC\_71 Homo sapiens cDNA IMAGE:5520909  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BMS42545  
 VERSION BMS42545.1 GI:18772149  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgs.nci.nih.gov/  
 1 (bases 1 to 1143)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rcmail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: LLM12185 row: k column: 22  
 High quality sequence stop: 666.  
 Location/Qualifiers

## FEATURES

source

1..1143  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5520909"  
 /issue\_type="leiomysarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1ib="NIH MGC 71"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; NotI;  
 site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.1 kb."

## ORIGIN

Query Match 21.0%; Score 563; DB 4; Length 1143;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-269;  
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 2039 CTGACACCTTTCAGCACTGAGAGTGAAGAGAGCGTGTAAATACAGAGACAGA 2158  
 37 CTGACACCTTTCAGCACTGAGAGTGAAGAGAGCGTGTAAATACAGAGACAGA 96  
 2159 CGTCAAGGTGTTTTCACAGCCCTGAGAGAGAGAGGACCAAGGCTCCGACAGTGTCT 2218  
 97 CGTCAAGGTGTTTTCACAGCCCTGAGAGAGAGAGGACCAAGGCTCCGACAGTGTCT 156  
 2219 GGGGTGACTCTTGTGAGAGCTTTTACCTCTGAGTGAACCTTCCAGAGCCCGGG 2278  
 157 GGGGTGACTCTTGTGAGAGCTTTTACCTCTGAGTGAACCTTCCAGAGCCCGGG 216  
 2279 GGGGTGACTCTTGTGAGAGCTTTTACCTCTGAGTGAACCTTCCAGAGCCCGGG 2338  
 217 GGGGTGACTCTTGTGAGAGCTTTTACCTCTGAGTGAACCTTCCAGAGCCCGGG 276  
 2339 GAGCAAGCTTCTGTAAATGAGAGAGAGTGTGTTTCAAGAAAGTTTCAAGAGAGA 2398  
 277 GAGCAAGCTTCTGTAAATGAGAGAGAGTGTGTTTCAAGAAAGTTTCAAGAGAGA 336  
 2399 ACATTTTCGAAATCTACAGTTAAGCAGAGAGTGTGTTTCAAGAAAGTTTCAAGAGAGA 2458  
 337 ACATTTTCGAAATCTACAGTTAAGCAGAGAGTGTGTTTCAAGAAAGTTTCAAGAGAGA 396

QY 2459 AGGCGAAGTTTATCAAAAAATTGTTTCAGAGAGAGGAGCATTAAGTTTACAGCTTACAG 2518  
 DB 397 AGGCGAAGTTTATCAAAAAATTGTTTCAGAGAGAGGAGCATTAAGTTTACAGCTTACAG 456  
 QY 2519 GAGGTACAAATATCTGCTGCTGGGAAACACACATTTTATCTATTTTATTTTAA 2578  
 DB 457 GAGGTACAAATATCTGCTGCTGGGAAACACACATTTTATCTATTTTATTTTAA 516  
 QY 2579 TAGGTTGGTCTTATCTTATTAAGATTAAATGTCACAACTGTACACAAATATA 2638  
 DB 517 TAGGTTGGTCTTATCTTATTAAGATTAAATGTCACAACTGTACACAAATATA 576  
 QY 2639 TAATTATTAATTACAAATTGAC 2661  
 DB 577 TAATTATTAATTACAAATTGAC 599

RESULT 14  
 BE407732 666 bp mRNA 1linear EST 21-JUL-2000  
 LOCUS 60189777F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3629682 5'  
 DEFINITION mRNA sequence.  
 ACCESSION BE407732  
 VERSION BE407732.1 GI:9344182  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 666)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LIML at: image.liml.gov  
 Plate: LINC316 row: J column: 19  
 High quality sequence stop: 640.  
 Location/Qualifiers  
 1..666  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3629682"  
 /issue\_type="choriocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_21"  
 /note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN  
 Query Match 20.9%; Score 560; DB 2; Length 666;  
 Best Local Similarity 99.8%; Freq. No. 5.2e-268;  
 Matches 610; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1156 AACCACTGCTGGAACATACCTCATCCAGATCCAGACAAGAGTGCAGTAAGAAGAT 1215  
 DB 1 AACCACTGCTGGAACATACCTCATCCAGATCCAGACAAGAGTGCAGTAAGAAGAT 60  
 QY 1216 GTGCAAAAGTATGATGCCAGGAATAAATCACTCAAGACATGCTGACGCCCAAGTCAG 1275  
 DB 61 GTGCAAAAGTATGATGCCAGGAATAAATCACTCAAGACATGCTGACGCCCAAGTCAG 120

QY 1276 CGGTCTTTTCTGATGAGAGAGAGTTTCAAGACCTGCTGAGCTGTACAGCTTGAC 1335  
 DB 121 CGGTCTTTTCTGATGAGAGAGAGTTTCAAGACCTGCTGAGCTGTACAGCTTGAC 180  
 QY 1336 AGTAGAGCTCTGACATTAAGCCAGCATTCGTCGTGTGCGGAGTGTCTTAAGTACAA 1395  
 DB 181 AGTAGAGCTCTGACATTAAGCCAGCATTCGTCGTGTGCGGAGTGTCTTAAGTACAA 240  
 QY 1396 AGGCAAGGCGGAGAGCTCCCACTGCGCAGAGCCGAGGAGGAGCCAGAGCCCAAG 1455  
 DB 241 AGGCAAGGCGGAGAGCTCCCACTGCGCAGAGCCGAGGAGGAGCCAGAGCCCAAG 300  
 QY 1456 GCCCTGGGAGATGACACCTTCACAGTCCGTCAAGCTTGAACAGACAGATCCAGATTACGT 1515  
 DB 301 GCCCTGGGAGATGACACCTTCACAGTCCGTCAAGCTTGAACAGACAGATCCAGATTACGT 360  
 QY 1516 TGCCCTTGCAAGAGAACCCAGCCCTGTGACATCTGCTTCAAGCCCATGCCCCAG 1575  
 DB 361 TGCCCTTGCAAGAGAACCCAGCCCTGTGACATCTGCTTCAAGCCCATGCCCCAG 420  
 QY 1576 AAGAGGAGCGGAGAGAGACCCGCGTGCCTTCAAGCAAGTGTGCGGTGCGCCGAG 1635  
 DB 421 AAGAGTGAAGCGGAGAGAGACCCGCGTGCCTTCAAGCAAGTGTGCGGTGCGCCGAG 480  
 QY 1636 CCTTCTGCACTGTACTGAGGCTGACACCCGAGCCGCTGTACAGCTGCTGCCCCG 1695  
 DB 481 CCTTCTGCACTGTACTGAGGCTGACACCCGAGCCGCTGTACAGCTGCTGCCCCG 540  
 QY 1696 TTTTGTAGCTCAACTGTGAGTGAACAAGTGTGACAGCGGTGCTGAACAACAACAGTAC 1755  
 DB 541 TTTTGTAGCTCAACTGTGAGTGAACAAGTGTGACAGCGGTGCTGAACAACAACAGTAC 600  
 QY 1756 GAGTCAGACAT 1766  
 DB 601 GAGTCAGACAT 611

RESULT 15  
 AL539797/c 1118 bp mRNA 1linear EST 24-MAR-2004  
 LOCUS AL539797 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
 DEFINITION CS0DF031YA23 3-PRIME, mRNA sequence.  
 ACCESSION AL539797  
 VERSION AL539797.3 GI:45715435  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 1118)  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:31264360.  
 CONTACT: Genoscope  
 Genoscope -- Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6792.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?e=CS0DF031YA12NP1&c=6792.r.

FEATURES  
 source  
 1..1118  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF031YA23"

```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10CM189 row: p column: 05
High quality sequence stop: 748.
location/Qualifiers
1. .934
/organism="Homo sapiens"
seq_method="dbm"

```

**FEATURES**  
**SOURCE**

/note="Organ: uterus; Vector: pOT7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adapter:  
 GGCGCGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 library."

```

II RT (Lite Technologies). Note: this is a NIH_MGC
Library."

```

QY	1944	GAACATTCCTGCTTCCGAGTTGCCAGTGGCCGTAACTCCCTCCTGACTGCTACTGAGG	2007
Db	699	GAACTTCCTGCTTCCGAGTTGCCAGTGGCCGTAACTCCCTCCTGACTGCTACTGAGG	640
QY	2004	CCGTACTCCGCGACTCAGGTGAAGCTCACACGCGCATGAAATTCATCATCTGTGA	2065
Db	639	CCGTACTCCGCGACTCAGGTGAAGCTCACACGCGCATGAAATTCATCATCTGTGA	580
QY	2064	ACAGCAAGGTTCAAAAACTAAGCTCAGAGGCCCTGAGCACTTTCAGCACTGAGGT	2122
Db	579	ACAGCAAGGTTCAAAAACTAAGCTCAGAGGCCCTGAGCACTTTCAGCACTGAGGT	520
QY	2124	GAAAGAGCGTGTTTTAAATACAGAGCAAGCAAGCAAGTGTGTTTCAACGCCCTT	2188
Db	519	GAAAGAGCGTGTTTTAAATACAGAGCAAGCAAGCAAGTGTGTTTCAACGCCCTT	460
QY	2184	GAGGGAAGGAGCGCAGAGGCTCCGACAGGTGCTCAGGGGTGACTGTTCTGTGAGCTTT	2242
Db	459	GAGGGAAGGAGCGCAGAGGCTCCGACAGGTGCTCAGGGGTGACTGTTCTGTGAGCTTT	401

RESULT	16
LOCUS	BF971838
DEFINITION	BF971838 934 bp mRNA linear EST 22-JAN-2001 602240424r1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4329076 5', mRNA sequence.
ACCESSION	BF971838
VERSION	BF971838
KEYWORDS	BF971838.1 GI:12339053
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 934) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
REFERENCE	Contact: Robert Strausberg, Ph.D.
AUTHORS	Email: cgabbs@mail.nih.gov
JOURNAL	Tissue Procurement: ATCC
COMMENT	cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.

QY	1208	GGGAGAGCTCCGCTGAGCTGACCTATTCAGATATCGGAGAGAAACATTCCTGCTCCGAGTTGACC	196
Db	1	GGGAGAGCTCCGCTGAGCTGACCTATTCAGATATCGGAGAGAAACATTCCTGCTCCGAGTTGACC	60
QY	1368	AGTGGCCGTAAACATCCCGTCTGACTGCTAATCGGGCCGTAACTGCGGACTCAGGTGAA	2022
Db	61	AGTGGCCGTAAACATCCCGTCTGACTGCTAATCGGGCCGTAACTGCGGACTCAGGTGAA	120
QY	2028	AGCTCACCACGCGCATGAAATTCATTCATCTGTGAAACAGCAAGGTTCAAAACTAAGC	208
Db	121	AGCTCACCACGCGCATGAAATTCATTCATCTGTGAAACAGCAAGGTTCAAAACTAAGC	180
QY	2088	ATCCAGAGGCCCTGAGAGCTTTTCAGACACTGGAAGTGAAGAGAGCGTGTTTTAAATAC	2144
Db	181	ATCCAGAGGCCCTGAGAGCTTTTCAGACACTGGAAGTGAAGAGAGCGTGTTTTAAATAC	240
QY	2148	AGAGACAAGCAGCTCAAGGTGTTTTCAAGCCCTTGAGGGAGAGGACGCAAGGCTCTCG	2207
Db	241	AGAGACAAGCAGCTCAAGGTGTTTTCAAGCCCTTGAGGGAGAGGACGCAAGGCTCTCG	300
QY	2208	ACAGGTGCTCGGGGGTGACTCTTCTGAGAGACTTTTAAACCTCTGAGAGAGCCCTCCCG	226
Db	301	ACAGGTGCTCGGGGGTGACTCTTCTGAGAGACTTTTAAACCTCTGAGAGAGCCCTCCCG	360
QY	2268	AGAGCCCCCGGGGGCCGAGCCCGCCTCTGTGTAGCGCTGAGGACAGGGCTCGTGTGACA	2327
Db	361	AGAGCCCCCGGGGGCCGAGCCCGCCTCTGTGTAGCGCTGAGGACAGGGCTCGTGTGACA	419
QY	2328	TCAGCACAAGAGAGAAAGCTTTCTGTAAACAATGCGCCCTCCCGCAGAGAGGCGAAGTTT	2387
Db	420	TCAGCACAAGAGAGAAAGCTTTCTGTAAACAATGCGCCCTCCCGCAGAGAGGCGAAGTTT	479
QY	2388	TGCTCTTTTGTACATTTTCCGAAACTCAGTTAAAGCAGAAAGTCTGTTTTCAGGAAAAAGT	2447
Db	480	TGCTCTTTTGTACATTTTCCGAAACTCAGTTAAAGCAGAAAGTCTGTTTTCAGGAAAAAGT	539
QY	2448	TTCAAGGAGAGAGGCGAAGTTTATCAAAAACATTTGTTTCAGAGAGAGGAGATTAAGTTT	2507
Db	540	TTCAAGGAGAGAGGCGAAGTTTATCAAAAACATTTGTTTCAGAGAGAGGAGATTAAGTTT	599
QY	2508	ACAGCCTTACAGAGCTTACAAATATCCGTGCTGGGAAA	2547
Db	600	ACAGCCTTACAGAGCTTACAAATATCTGCTGCTGGGAAA	639

RESULT 17  
 CA916737 729 bp mRNA linear EST 14-MAR-2003  
 LOCUS 301SPRIM2-5A human lymphocyte Matchmaker cDNA library Homo sapiens  
 DEFINITION cDNA clone L5 5' similar to CHFR RF; cell cycle check point, mRNA  
 sequence.  
 CA916737  
 ACCESSION CA916737.1 GI:28951974  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 729)  
 Scarafla, L.E., Stouffer, S.D. and Swinney, D.C.  
 Identification of Ring Finger proteins that interact with Ubch5a,  
 an ubiquitin-conjugating enzyme  
 Unpublished (2002)  
 JOURNAL Contact: Scarafla LE  
 COMMENT Enzymology  
 Roche Bioscience  
 3401 Hillview Ave, S3-1, Palo Alto, CA 94304-1397, USA  
 Tel: 650 354 7997  
 Fax: 650 354 7554  
 Email: liliana.scarafila@roche.com  
 Seq primer: matchmaker 5' AD (ctatcgatgatgaagaatccccaccaccc)  
 POLYA=No.

FEATURES  
 source Location/Qualifiers  
 1..729  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="L5"  
 /cell\_type="B cell"  
 /cell\_line="EBV-transformed human peripheral blood  
 lymphocyte; B cell population; IG (+)"  
 /lab\_host="yeast/E.coli"  
 /clone\_lib="human lymphocyte Matchmaker cDNA library"  
 /note="Vector: PACT; Site\_1: XhoI; Site\_2: XhoI;  
 Matchmaker yeast two-hybrid system from Clontech; PACT has  
 GAL4 AD under ADHI promoter. This oligo-dr primed library  
 was screened with human Ubch5a as bait, to obtain  
 interacting proteins."

ORIGIN  
 Query Match 19.1%; Score 513; DB 6; Length 729;  
 Best Local Similarity 99.5%; Pred. No. 1.7e-244;  
 Matches 663; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 560 CCGATCCCGGGTCCCTCGCTGCGCCGCGCACTGAGTGTGCTTGAAGAACACAGC 619  
 |||||  
 DB 13 CCGATCCCGGGTCCCTCGCTGCGCCGCGCACTGAGTGTGCTTGAAGAACACAGC 72  
 |||||

QY 620 CATCAACATGACGTGACAGCTTCCCAACAGCTCGGCTTCCACAGAGCCCTTC 679  
 |||||  
 DB 73 CATCAACATGACGTGACAGCTTCCCAACAGCTCGGCTTCCACAGAGCCCTTC 132  
 |||||

QY 680 CTGCAAGGCGAGAGCGTTCCTCAAGTGTGGGTCTGGGGGTGTGGCATCTCCCTTAAG 739  
 |||||  
 DB 133 CTGCAAGGCGAGAGCGTTCCTCAAGTGTGGGTCTGGGGGTGTGGCATCTCCCTTAAG 192  
 |||||

QY 740 GAAAGTGTCCCTGTGCGCAAGTGAAGTCTCCAGCTTGGCTTCCACAGCA 799  
 |||||  
 DB 193 GAAAGTGTCCCTGTGCGCAAGTGAAGTCTCCAGCTTGGCTTCCACAGCA 252  
 |||||

QY 800 GAAAGCTGGTCTTGTGCGTGAAGCCCAAGATGAGAGATTGGAGCCGGTA 859  
 |||||  
 DB 253 GAAAGCTGGTCTTGTGCGTGAAGCCCAAGATGAGAGATTGGAGCCGGTA 312  
 |||||

QY 860 AGAAGAAAATGAGAGAGATGGGACCTTGACCTGAACGGGCGATTGTGTGCAAAAC 919  
 |||||  
 DB 313 AGAAGAAAATGAGAGAGATGGGACCTTGACCTGAACGGGCGATTGTGTGCAAAAC 372  
 |||||

QY 920 CGCGTAGAAATGCCAAACCCCTCCAGAGAGCTCAGAGACGGCTGGGAAGCCAGACA 979  
 |||||  
 DB 373 CGCGTAGAAATGCCAAACCCCTCCAGAGAGCTCAGAGACGGCTGGGAAGCCAGACA 432  
 |||||

QY 980 AGATGAGAGAGAGCTGATCATCATCTGTCAGAGACCTGCTGACAGACTGGTAGTT 1039  
 |||||  
 DB 433 AGATGAGAGAGAGCTGATCATCATCTGTCAGAGACCTGCTGACAGACTGGTAGTT 492  
 |||||

QY 1040 TCGAGCCCTGATGCAACACGTTCTGGCGGCTTGTCTACTCGGGCTGGATGAGACGCTCGT 1099  
 |||||  
 DB 493 TCGAGCCCTGATGCAACACGTTCTGGCGGCTTGTCTACTCGGGCTGGATGAGACGCTCGT 552  
 |||||

QY 1100 CCCTGTGCTTACCTCGCGCTGCTCCCGTGGAGCGGATCTGTAAAACCATCTCTCAACA 1159  
 |||||  
 DB 553 CCCTGTGCTTACCTCGCGCTGCTCCCGTGGAGCGGATCTGTAAAACCATCTCTCAACA 612  
 |||||

QY 1160 ACCCTGTGAAACATACCTCATTCAGACATCCAGACAAAGATGCAAGTAAGATGTGC 1219  
 |||||  
 DB 613 ACCCTGTGAAACATACCTCATTCAGACATCCAGACAAAGATGCAAGTAAGATGTGC 672  
 |||||

QY 1220 AAAATA 1225  
 |||||  
 DB 673 AAAATA 678  
 |||||

RESULT 18  
 BE280165 690 bp mRNA linear EST 13-JUL-2000  
 LOCUS 601158370F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3504928 5',  
 DEFINITION mRNA sequence.  
 BE280165  
 ACCESSION BE280165.1 GI:9155072  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 690)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgarbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA library Preparation: Ling Hong/Rubin laboratory  
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNL at: image.ilnl.gov  
 Plate: DLM180 row: 1 column: 17  
 High quality sequence stop: 626.

FEATURES  
 source Location/Qualifiers  
 1..690  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3504928"  
 /issue\_type="choriocarcinoma"  
 /lab\_host="PH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 21"  
 /note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACTAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN  
 Query Match 18.2%; Score 488; DB 2; Length 690;  
 Best Local Similarity 99.8%; Pred. No. 5.3e-232;











Query Match	17.0%;	Score 456;	DB 4;	Length 1096;
Best Local Similarity	99.8%;	Pred. No. 5.3e-216;		

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:555112"  
/tissue_type="retinoblastoma"
```

/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_67"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 Kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 16.9%; Score 452; DB 4; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 5.3e-214;  
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 709 GGGTCGCGGGGTGGGCAATCTCCCTAAGAGAGTGTCTCTGTGGCAATGATGA 768  
DB 150 GGGTCGCGGGGTGGGCAATCTCCCTAAGAGAGTGTCTCTGTGGCAATGATGA 209  
QY 769 GTCTCAAGCTTTGCTCTGAGCTCTCCAGACAGAAAGCTGCTCTTTTGTCTGGAA 828  
DB 210 GTCTCAAGCTTTGCTCTGAGCTCTCCAGACAGAAAGCTGCTCTTTTGTCTGGAA 269  
QY 829 CCCAGATCAGAGAGATTTGAGCCCGTGAAGAAATGAGAGATGGGGACCTT 888  
DB 270 CCCAGATCAGAGAGATTTGAGCCCGTGAAGAAATGAGAGATGGGGACCTT 329  
QY 889 GACCTGAACGGGCAAGTTGTTGTGCAACAACCGCTAGAATGCCAAACCGTCAAG 948  
DB 330 GACCTGAACGGGCAAGTTGTTGTGCAACAACCGCTAGAATGCCAAACCGTCAAG 389  
QY 949 GACCTGAACGGGCAAGTTGTTGTGCAACAACCGCTAGAATGCCAAACCGTCAAG 1008  
DB 390 GACCTGAACGGGCAAGTTGTTGTGCAACAACCGCTAGAATGCCAAACCGTCAAG 449  
QY 1009 TGCAGAGACCTGCTGACAGCTGCTGAGTTTTCAGCCCTGACAGCAACGTTCTGCGG 1068  
DB 450 TGCAGAGACCTGCTGACAGCTGCTGAGTTTTCAGCCCTGACAGCAACGTTCTGCGG 509  
QY 1069 GCTTGTACTCGGGCTGAGTGAAGCGCTCGTCCCTGTGTCCTACTGCGCTGCGG 1128  
DB 510 GCTTGTACTCGGGCTGAGTGAAGCGCTCGTCCCTGTGTCCTACTGCGCTGCGG 569  
QY 1129 GAGCGGATCTGTAAAAACCAATCTCAACA 1160  
DB 570 GAGCGGATCTGTAAAAACCAATCTCAACA 601

RESULT 28  
B0936072 953 bp mRNA linear EST 21-AUG-2002  
LOCUS AGENCOURT 8779684 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6371968  
DEFINITION 5', mRNA sequence.  
ACCESSION B0936072  
VERSION B0936072.1 GI:22351455  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 953)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF/Gazdar  
cDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNC2546 row: p column: 17  
High quality sequence stop: 630.  
Location/Qualifiers

## source

1. .953  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:6371968"  
/clone\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Lung; Vector: pOT7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 16.8%; Score 450; DB 5; Length 953;  
Best Local Similarity 100.0%; Pred. No. 5.3e-213;  
Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GCGGACCCGTTTCGGGTTTCGGCGCGGGGATGTGAATCCGATGAGCGCCCGA 104  
DB 1 GCGGACCCGTTTCGGGTTTCGGCGCGGGGATGTGAATCCGATGAGCGCCCGA 60  
QY 105 GGAAGGCAAGCAAGTCCGCGCGCGCGGAGCCCTGGGGAAGGCTCTGCGGCGCGGA 164  
DB 61 GGAAGGCAAGCAAGTCCGCGCGCGCGGAGCCCTGGGGAAGGCTCTGCGGCGCGGA 120  
QY 165 GGAAGGCAAGCAAGTCCGCGCGCGCGGAGCCCTGGGGAAGGCTCTGCGGCGCGGA 224  
DB 121 GGAAGGCAAGCAAGTCCGCGCGCGCGGAGCCCTGGGGAAGGCTCTGCGGCGCGGA 180  
QY 225 TTGCGACCTTTCTTCCCAAGCAATGCTGTCTGAGATCACTGTAATTTAGT 284  
DB 181 TTGCGACCTTTCTTCCCAAGCAATGCTGTCTGAGATCACTGTAATTTAGT 240  
QY 285 GATGAAAAATCAGGTCAAGTGAACATGGAATACCAACCAAGTGAACATGATTTAA 344  
DB 241 GATGAAAAATCAGGTCAAGTGAACATGGAATACCAACCAAGTGAACATGATTTAA 300  
QY 345 CAAGCTGAAGTTGTTAAGAGAGACATGCCCTTACAGACTGGGATGTCATCTACT 404  
DB 301 CAAGCTGAAGTTGTTAAGAGAGACATGCCCTTACAGACTGGGATGTCATCTACT 360  
QY 405 GGTGTACAGAAAGATGAACCGGAACAACGTCGATACCTTATGAATCTTTAAGTA 464  
DB 361 GGTGTACAGAAAGATGAACCGGAACAACGTCGATACCTTATGAATCTTTAAGTA 420  
QY 465 AAGCAAGGATGACACAGAATCTTTGA 494  
DB 421 AAGCAAGGATGACACAGAATCTTTGA 450

RESULT 29  
B0924736 938 bp mRNA linear EST 20-AUG-2002  
LOCUS AGENCOURT 8853777 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6374110  
DEFINITION 5', mRNA sequence.  
ACCESSION B0924736  
VERSION B0924736.1 GI:22339767  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 938)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF/Gazdar  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LNCM2552 row: 1 column: 23  
 High quality sequence scop: 629.  
 Location/Qualifiers

## FEATURES

source

```
1..938
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6374110"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_18"
/Note="Origin: lung; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

## ORIGIN

Query Match 16.7%; Score 447; DB 5; Length 938;

Best Local Similarity 100.0%; Pred. No. 1.7e-211; Indels 0; Gaps 0;

Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 48 CAGCCGGTTCGGGTCGGCGCGGGGCGGAGTGTGAATCCGATGAGCGGCCGAGGA 107
DB 1 CAGCCGGTTCGGGTCGGCGCGGGGCGGAGTGTGAATCCGATGAGCGGCCGAGGA 60
QY 108 AGGCAAGCAGTGGCCCGCCCGCAGCCCTGGGAGCGGCTTCCTGCTGGGCGCGGAGA 167
DB 61 AGGCAAGCAGTGGCCCGCCCGCAGCCCTGGGAGCGGCTTCCTGCTGGGCGCGGAGA 120
QY 168 GGGCGGCGCGAGCTCTCTGAGAGAGGGAGTGGACCATCGGGCGGAGACGAGTTG 227
DB 121 GGGCGGCGCGAGCTCTCTGAGAGAGGGAGTGGACCATCGGGCGGAGACGAGTTG 180
QY 228 CGACCTTCTCTCCCGAGCAATTAATCTGTCTCTGAGATCACTGTGAATTTAGTGA 287
DB 181 CGACCTTCTCTCCCGAGCAATTAATCTGTCTCTGAGATCACTGTGAATTTAGTGA 240
QY 288 TGAATAATCAGTCAAGTGAACACTGGAAGATACACAGCAGTGGAACTGATTACAA 347
DB 241 TGAATAATCAGTCAAGTGAACACTGGAAGATACACAGCAGTGGAACTGATTACAA 300
QY 348 GGTGAAGGTGTTAAGAGAGACATGCGCTTACAGACTGGGGAGTGTATCTACTTGT 407
DB 301 GGTGAAGGTGTTAAGAGAGACATGCGCTTACAGACTGGGGAGTGTATCTACTTGT 360
QY 408 GTACAGGAAGATGAACCGGAACACAACTGGGATACCTCTATGAATCTTTAAGTAAA 467
DB 361 GTACAGGAAGATGAACCGGAACACAACTGGGATACCTCTATGAATCTTTAAGTAAA 420
QY 468 GCAAGGCATGACACAAATCTTTTGA 494
DB 421 GCAAGGCATGACACAAATCTTTTGA 447
```

## RESULT 30

BX486289 535 bp mRNA linear EST 04-SEP-2003  
 LOCUS BX486289  
 DEFINITION DKFZp686P03250.r1.686 (synonym: hlc3) Homo sapiens cDNA clone  
 ACCESSION BX486289  
 VERSION BX486289.1 GI:31949829  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 535)  
 AUTHORS Bloecher,H., Boecker,M., Mewes,H.W., Well,B., Amid,C., Oanger,A.,  
 Fobg,G., Han,M. and Wiemann,S.  
 EST (Bioecker,H., Boecker,M., Mewes,H.W., Well,B., Amid,C., et al.)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: MIPS

Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
 Braunschweig/Germany) within the cDNA sequencing consortium of the  
 German Genome Project.

No 81 sequence available.  
 This clone (DKFZp686P03250) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source

```
1..535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686P03250"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="686 (synonym: hlc3)"
/Note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIb;
cDNA-collection"
```

## ORIGIN

Query Match 16.3%; Score 436; DB 5; Length 535;

Best Local Similarity 100.0%; Pred. No. 5.3e-206; Indels 0; Gaps 0;

Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 59 GGGTTCGGCGCGGGGCGGAGTGTGAATCCGATGAGCGGCCGAGGAGCAAGCACT 118
DB 2 GGGTTCGGCGCGGGGCGGAGTGTGAATCCGATGAGCGGCCGAGGAGCAAGCACT 61
QY 119 GCGCCCGCGCGAGCCCTCGGGGAGCGGCTCTGCGCTCTGGGCGCGGAGGCGGAGCCGC 178
DB 62 GCGCCCGCGCGAGCCCTCGGGGAGCGGCTCTGCGCTCTGGGCGCGGAGGCGGAGCCGC 121
QY 179 AGTCTCTCTGAGAGAGGGGAGTGGACCATCGGGCGGAGAGAGGTTCCACTTCT 238
DB 122 AGTCTCTCTGAGAGAGGGGAGTGGACCATCGGGCGGAGAGAGGTTCCACTTCT 181
QY 239 TCCCGAGCAATTAATCTGTCTCTGAGATCACTGTGAATTTAGTGAATTAATCAG 298
DB 182 TCCCGAGCAATTAATCTGTCTCTGAGATCACTGTGAATTTAGTGAATTAATCAG 241
QY 299 GTCAAGTGAACACTGGAAGATACACAGCAGTGGAACTGATTAAACAAGCTGAAGTTG 358
DB 242 GTCAAGTGAACACTGGAAGATACACAGCAGTGGAACTGATTAAACAAGCTGAAGTTG 301
QY 359 TTAAGAGAGAGCATGCCCTTACAGACTGGGGAGTGTATCTACTGTTGTACAGAGAGA 418
DB 302 TTAAGAGAGAGCATGCCCTTACAGACTGGGGAGTGTATCTACTGTTGTACAGAGAGA 361
QY 419 ATGAACCGGAACACAGCTGGATACCTCTATGAATCTTTAAGTAAAAAGCAAGCATGA 478
DB 362 ATGAACCGGAACACAGCTGGATACCTCTATGAATCTTTAAGTAAAAAGCAAGCATGA 421
QY 479 CACAAGATCTTTTGA 494
DB 422 CACAAGATCTTTTGA 437
```

## RESULT 31



EX486564	557 bp	mRNA	linear	EST 04-SEP-2003
LOCUS				
DEFINITION	DKRZP0666H04252_r1 666 (synonym: hlcc3) Homo sapiens cDNA clone DKRZP0666H04252_5, mRNA sequence.			
ACCESSION	BX486564			
VERSION	BX486564			
KEYWORDS	BX486564.1 GI:31950352			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Bukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 557)			
REFERENCE	Bahr, A., Lauber, J., Mewes, H.W., Well, B., Amlid, C., Osanger, A., Fobo, G., Han, M., and Wiemann, S.			
AUTHORS	EST (Bahr, A., Lauber, J., Mewes, H.W., Well, B., et al.)			
TITLE	Unpublished (2003)			
JOURNAL	Contact: MIPS			
COMMENT				

Ingolstaedter Lander 1, D-85764 Neuburg, Germany. This is the 5' sequence of the clone insert from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de, consortium of the German Genome Project. No *si* sequence available. This clone (DKFZp686H04252) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg; GEMNAV; Email: clone@rzpd.de.

**FEATURES**  
**source**

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDK2p686H0452"
/db_host="dnalib"
/clone_lib="666 (synonym: hlcc3)"
/name="Vector: pprIDEX2; site_1: s111; site_2: s111b;
cdna_collection"

```

## ORIGIN

Query Match	16.3%;	Score 436;	DB 5;	Length 557;
Best Local Similarity	100.0%;	Pred. NO. 5.3e-206;		
Matches 436;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

Qy 119 CGCGCGCGCGAGCCCTGSGGACGGTCTCTTCGCTCTGGGCGCGAGAGAGGGCGAGCCGC 178  
 Db 62 CGCGCGCGCGAGCCCTGSGGACGGTCTCTTCGCTCTGGGCGCGAGAGAGGGCGAGCCGC 121

179 ACGTCTCTCAGGGAAGCGGAGTGGACCAATCGGGCGGAGAGAGTTCGACCTTTCCT 238  
 |||  
 122 ACGTCTCTCAGGGAAGCGGAGTGGACCAATCGGGCGGAGAGAGTTCGACCTTTCCT 181

235 TCCGAGCAATAAACGGCTCTGGAGATCACTGTGAATGTAGTGAATGAAAAATCAG 298  
Db 182 TCCGAGCAATAAACTGCTGTGAGATCACTGTGAATGTAGTGAATGAAAAATCAG 241

242 GTCTAGCTGACCTGGAGATACCAACACCACTGGAGCACTGATTACCAAGCTGAGGTTG 301

Db 302 TTAAGAAGCAGACATGCCCTTACAGACTGGGATGCATCTACTTGGTACAGGAAGA 361

Db 362 ATGACCGAACACAACTGTCATGAACTTTAAGTAAAGCAAGCATGA 421

QY	479	CACAAGAATCCTTTGA	494
Db	422	CACAAGAATCCTTTGA	437

RESULT 32	
CB989147	
LOCUS	811 bp mRNA, linear EST 01-MAY-2003
DEFINITION	AGENCOURT 13893036 NIH MGC 147 Homo sapiens cDNA clone IMAGE:30342520 5', mRNA sequence.

ACCESSION	CB989147
VERSION	CB989147.1
	GI:30283667

NEIMONDS	ESI.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 811)  
NIN-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC).  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Dr. Stefan Hansson  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: NDAM375 row: 1 column: 17  
High quality sequence stop: 476.

## FEATURES

### Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30342520"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_1ib="NIH MGC 147"
/note="Organ: Placenta; Vector: pBluescript; Site 1:
5'-tho1; Site 2: BamH1 Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTTTV-3', size-selected for average
insert size 2.3 kb and normalized to ROI 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carinci, in
preparation). Library constructed by M. Brownstein
(NIMH/NIHGR1, National Institutes of Health). Note: This is
a NIH MGC library."

```

**ORIGIN**

Query Match	16.2%;	Score 435;	DB 6;	Length 811;
Best Local Similarity	100.0%;	Pred. No. 1.7e-205;		
Matches 435;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

0y      60 GGTTCGCGCCCGGGCGGGGATGTGAATCCCGATGAGCGGCCGAGAAAGCCAAAGCATTC 119
      |||
Db      27 GGTTCGCGCCCGGGCGGGGATGTGAATCCCGATGAGCGGCCGAGAAAGCCAAAGCATTC 86

```

**Oy** 120 GCGGCGCCGCAGCCCTGGGGACGGCTCCTGCGCTCTGGGCGGGAGGAGGGCGAAGCCCCA 179  
**Dd** 87 GCCGCCGCCGCAGCCCTTGGGGACGGCTCTGCGCTCTGGGCGGGAGGAGGGCGAAGCCCCA 146

180 CCGTTCCTCGAGGAGCGGAGTGGACCAATCGGGCGGAGACGAGGTTGCGACTTCTCTT 206

207 CCCGACATAACTGGTCTCTGGAGATCACTGAGATTGTAGTGGATGAAAAATCAGG 266

QY 300 TCAGGTGACACTGGAAGATACCAAGCAACGTGGAACGTGATTAACAAGCTGAAGTTCT 359  
 DB 267 TCAGGTGACACTGGAAGATACCAAGCAACGTGGAACGTGATTAACAAGCTGAAGTTCT 326  
 QY 360 TAGAAGACGACATGCGCTTTACAGCTGGGATGTCATCTACTTGGTGTACAGGAAGA 419  
 DB 327 TAGAAGACGACATGCGCTTTACAGCTGGGATGTCATCTACTTGGTGTACAGGAAGA 386  
 QY 420 TGAACCGGAACAACAAGTGCATACCTCTATGATATCTTTAAGTGAAGCAAGGATGAC 479  
 DB 387 TGAACCGGAACAACAAGTGCATACCTCTATGATATCTTTAAGTGAAGCAAGGATGAC 446  
 QY 480 ACAAGATCTTTGA 494  
 DB 447 ACAAGATCTTTGA 461

RESULT 33  
 CB961597 806 bp mRNA linear EST 29-APR-2003  
 LOCUS AGENCOURT\_13892092 NIH\_MGC\_148 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30337709 5', mRNA sequence.  
 CB961597  
 VERSION CB961597.1 GI:30217714  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 806)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaab-r@mail.nih.gov  
 Tissue Procurement: Dr. Stefan Hansson  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
 and advice from Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: NDAM363 row: d column: 06  
 High quality sequence stop: 542.  
 Location/Qualifiers  
 1..806  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30337709"  
 /tissue\_type="pre-ecclampic placenta"  
 /lab\_host="DH10B TONa"  
 /clone\_id="NIH\_MGC\_148"  
 /note="Organ: Placenta; Vector: pBluescriptR; Site\_1:  
 att-XhoI; Site\_2: BamH; Library is oligo-dT primed and  
 directionally cloned using primer  
 5'-TTTTTATTTTATTTVN-3', size-selected for average insert  
 size 2.3 kb and normalized to ROP 5. This is a primary  
 library enriched for full-length clones and constructed  
 using the Cap-trapper method (Carninci, in preparation).  
 Library constructed by M. Brownstein (NIH/NHGRI,  
 National Institutes of Health). Note: this is a NIH\_MGC  
 library."

ORIGIN  
 Query Match 16.2%; Score 433; DB 6; Length 806;  
 Best Local Similarity 100.0%; Pred.No.1.7e-204;  
 Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 62 TTCGGCGCGGCGGCGGATGTGATCCGATGGAGCGGCCGAGGAAGCAAGCAAGTCCG 121  
 DB 29 TTCGGCGCGGCGGCGGATGTGATCCGATGGAGCGGCCGAGGAAGCAAGCAAGTCCG 88

QY 122 GCGCCGCCGAGCCTCTGGGAGAGCGCTCTGCGCTCTGGGCGCGGAGGAGCGGCCGACG 181  
 DB 89 GCGCCGCCGAGCCTCTGGGAGAGCGCTCTGCGCTCTGGGCGCGGAGGAGCGGCCGACG 148  
 QY 182 TCCTCTGAGGAGGAGGAGGAGTGAACCATGGGCGGAGAGAGTTGGACCTTTCCTCC 241  
 DB 149 TCCTCTGAGGAGGAGGAGGAGTGAACCATGGGCGGAGAGAGTTGGACCTTTCCTCC 208  
 QY 242 CCAGCAATTAACCTGCTCTGAGAGATCACTGTAGATTTGATGATGATGATGATGATGAT 301  
 DB 209 CCAGCAATTAACCTGCTCTGAGAGATCACTGTAGATTTGATGATGATGATGATGATGAT 268  
 QY 302 AGGTGACACTGGAAGATACCAAGCAACCATGGAACAGTATTAACAAGCTGAAGTTGTTA 361  
 DB 269 AGGTGACACTGGAAGATACCAAGCAACCATGGAACAGTATTAACAAGCTGAAGTTGTTA 328  
 QY 362 AGAAGCAGACATGCGCTTTACAGACTGGGAGATGTCATCTACTTGGTGTACAGGAAGATG 421  
 DB 329 AGAAGCAGACATGCGCTTTACAGACTGGGAGATGTCATCTACTTGGTGTACAGGAAGATG 388  
 QY 422 AACCGGAACAACAAGTGCATACCTCTATGATCTTTAAGTGAAGCAAGGATGACAC 481  
 DB 389 AACCGGAACAACAAGTGCATACCTCTATGATCTTTAAGTGAAGCAAGGATGACAC 448  
 QY 482 AAGAATCTTTGA 494  
 DB 449 AAGAATCTTTGA 461

RESULT 34  
 AU154212/c 518 bp mRNA linear EST 05-AUG-2002  
 LOCUS AU154212 NT2RP4 Homo sapiens cDNA clone NT2RP400455 3', mRNA  
 DEFINITION AU154212  
 VERSION AU154212  
 KEYWORDS AU154212.1 GI:11015733  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 518)  
 Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,  
 Ozawa,M., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y.  
 and Isogai,T.  
 HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,  
 Saito,K., Yamamoto,J., Ozawa,M., Nishikawa,T., Nakamura,Y.,  
 Nagai,T., Sugano,S., Masuno,Y., Isogai,T.)  
 Unpublished (2000)  
 Contact: Takao Isogai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute.  
 Location/Qualifiers  
 1..518  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2RP400455"  
 /cell\_type="teratocarcinoma"  
 /cell\_line="NT2"  
 /clone\_id="NT2RP4"  
 /note="Vector: pMT186FL3; mRNA from NT2 neuronal precursor  
 cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Query Match 15.9%; Score 426; DB 1; Length 518;  
 Best Local Similarity 99.8%; Pred. No. 5.3e-201;  
 Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2185 AAGGAAAGGACGCGAGGGTCTCCGACAGGTCTGTGGGTGACTCTTCTGTGAGCTTTT 2244  
 DB 503 AAGGAAAGGACGCGAGGGTCTCCGACAGGTCTGTGGGTGACTCTTCTGTGAGCTTTT 444  
 QY 2245 ACCCTTGTAGTGAACCTTCCCGAGAGCCCGGGGGCCCGAGCCGCCCTTCTGTGTGAGC 2304  
 DB 443 ACCCTTGTAGTGAACCTTCCCGAGAGCCCGGGGGCCCGAGCCGCCCTTCTGTGTGAGC 384  
 QY 2305 GCTGGCAGGGCTCGTGTGTGATGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2364  
 DB 383 GCTGGCAGGGCTCGTGTGTGATGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324  
 QY 2365 CGTCCCGCGAGAGGGGAGATTGCTCTTTGTGACATTTTCCGAAACTACAGTTAAAGC 2424  
 DB 323 CGTCCCGCGAGAGGGGAGATTGCTCTTTGTGACATTTTCCGAAACTACAGTTAAAGC 264  
 QY 2425 AGAAGTCTGTTTCAAGGAAAGTTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2484  
 DB 263 AAAAGTCTGTTTCAAGGAAAGTTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 204  
 QY 2485 TCAG 2544  
 DB 203 TCAG 144  
 QY 2545 AAAACACAGCATTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 2604  
 DB 143 AAAACACAGCATTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 84  
 QY 2605 GATTAAATGTCAAACTGTAGCACAATAATATTAATTAATTAATTAATTAATTAATTA 2661  
 DB 83 GATTAAATGTCAAACTGTAGCACAATAATATTAATTAATTAATTAATTAATTAATTAAT 27

RESULT 35  
 AUI38310 687 bp mRNA linear EST 02-AUG-2002  
 LOCUS AUI38310 PLACE1 Homo sapiens cDNA clone PLACE1008304 5', mRNA  
 DEFINITION AUI38310  
 ACCESSION AUI38310  
 VERSION AUI38310.1 GI:10999831  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 687)  
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,  
 Yamamoto,J., Wakanatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and  
 Isegai,T.  
 TITLE HRI human cDNA project  
 JOURNAL Unpublished (2000)  
 COMMENT  
 CONTACT: Takao Isegai  
 Genomics Laboratory  
 Helix Research Institute  
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 Tel: 81-438-52-3975  
 Fax: 81-438-52-3986  
 Email: genomics@hri.co.jp  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute  
 FEATURES  
 source  
 1..687  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACE1008304"

Query Match 15.7%; Score 421; DB 1; Length 687;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-198;  
 Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 720 TGGTGGCATCTTCCCTTAAAGAGAGTGTCCCTCTGTGGCAATGATGAAGTTCAGACTT 779  
 DB 1 TGGTGGCATCTTCCCTTAAAGAGAGTGTCCCTCTGTGGCAATGATGAAGTTCAGACTT 60  
 QY 780 TGCCTCAGCTTCCCGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839  
 DB 61 TGCCTCAGCTTCCCGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
 QY 840 GGAGAGATTGAG 899  
 DB 121 GGAGAGATTGAG 180  
 QY 900 GCAATTGTTGTGTGACAAACCGGTAGAAATGCCAAACCGTCCACAGAGAGTCAAGAC 959  
 DB 181 GCAATTGTTGTGTGACAAACCGGTAGAAATGCCAAACCGTCCACAGAGAGTCAAGAC 240  
 QY 960 AGCGGCTGGAG 1019  
 DB 241 AGCGGCTGGAG 300  
 QY 1020 GCTGCACAGACTGCTGAGAGTTTGCAGCCCTGACATGACAGTCTGCGGCTTGTCTATC 1079  
 DB 301 GCTGCACAGACTGCTGAGAGTTTGCAGCCCTGACATGACAGTCTGCGGCTTGTCTATC 360  
 QY 1080 GGGCTGATGAGAGCGCTGCTCTGTGCTTACTGCGGCTGCTGCGGCTGCGGAGATCTG 1139  
 DB 361 GGGCTGATGAGAGCGCTGCTCTGTGCTTACTGCGGCTGCTGCGGCTGCGGAGATCTG 420

RESULT 36  
 CN428707 590 bp mRNA linear EST 16-MAY-2004  
 LOCUS CN428707  
 DEFINITION 17000600232773 GRN\_PRENEM Homo sapiens cDNA 5', mRNA sequence.  
 ACCESSION CN428707  
 VERSION CN428707.1 GI:47416301  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 590)  
 Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,  
 Li,Y., Xu,C., Pang,R., Guegler,K., Rao,M.S., Mandalam,R.,  
 Lebkowski,J. and Stanton,L.W.  
 TITLE Transcriptional characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 JOURNAL Nat. Biotechnol. 22 (6): 707-716 (2004)  
 COMMENT  
 CONTACT: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 590 Std Error: 0.00.  
 FEATURES  
 source  
 1..590  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"

/tissue\_type="embryonic stem cell, retinoic acid and  
mitogen-treated hES cell line H7"  
/clone\_lib="GRN\_PRENED"  
/note="oligo dt primed, full-length enriched cDNA library  
from hES cell line H7 (p29) maintained in feeder-free  
conditions. Embryoid bodies were generated in the presence  
of all-trans retinoic acid and mitogens."

## ORIGIN

Query Match 15.7%; Score 420; DB 7; Length 590;  
Best Local Similarity 100.0%; Pred. No. 5.3e-198;  
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2242 TTATCCCTCTGAGTGAACCTTCCCGAGAGCCCGGAGCCGCTCTGCTG 2301  
100 TTATCCCTCTGAGTGAACCTTCCCGAGAGCCCGGAGCCGCTCTGCTG 159  
2302 AGCGCTGGGCGAGGCTCGTGTGGCATCAGCAGAGAGCAAGCTTCTGTAACATGC 2361  
160 AGCGCTGGGCGAGGCTCGTGTGGCATCAGCAGAGAGCAAGCTTCTGTAACATGC 219  
2362 GGCCTGCTCCGCGAGAGGCGAGGCTTGTCTTTGTGATTTTCGAAATACAGTTAA 2421  
220 GGCCTGCTCCGCGAGAGGCGAGGCTTGTCTTTGTGATTTTCGAAATACAGTTAA 279  
2422 AGCAGAGTCTGTTTTCAGGAAAAGTTTCAAGGAGAGAGGCAAGTTTACAAAACATT 2481  
280 AGCAGAGTCTGTTTTCAGGAAAAGTTTCAAGGAGAGAGGCAAGTTTACAAAACATT 339  
2482 GTTTCAGGAGAGGAGCATTAAGTTTACAGCCGACAGAGCAATATCTGCTGCT 2541  
340 GTTTCAGGAGAGGAGCATTAAGTTTACAGCCGACAGAGCAATATCTGCTGCT 399  
2542 GGGAAAACCAAGCATTTTATCTATTTTATTTTAAATAGTTGTGCTTATCTCTTAA 2601  
400 GGGAAAACCAAGCATTTTATCTATTTTATTTTAAATAGTTGTGCTTATCTCTTAA 459  
2602 TTAAGTTTAAATGTCAAACTGTAGCAAAATTAATTAATTAATTAATTAATTAAT 2661  
460 TTAAGTTTAAATGTCAAACTGTAGCAAAATTAATTAATTAATTAATTAATTAATTA 519

## RESULT 37

BO876252 876 bp mRNA linear EST 16-AUG-2002  
LOCUS AGENCOURT 8681666 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6376937  
DEFINITION 5', mRNA sequence.

ACCESSION BO876252  
VERSION BO876252.1 GI:22268258  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgi.mgi.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: c9abbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNI at:  
http://image.jnl.gov  
Plate: LHC2559 row: 0 column: 18  
High quality sequence stop: 678.  
Location/Qualifiers

## FEATURES

source  
1..876  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="IMAGE:6376937"  
/tissue\_type="normal pigmented retinal epithelium"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 43"  
/note="Organ: eye; Vector: pORF7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC library. |"

## ORIGIN

Query Match 15.7%; Score 420; DB 5; Length 876;  
Best Local Similarity 99.5%; Pred. No. 5.3e-198;  
Matches 660; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

1165 GTGAGAGATACCTTCATCAGATTCAGACAAAGTGCAGTGAAGATGTCAAGT 1224  
1 GTGAGAGATACCTTCATCAGATTCAGACAAAGTGCAGTGAAGATGTCAAGT 60  
1225 ATGAGTGCAGAGATTAATAATCACTCAAGACATGCTGACGCCCAAGTGAAGCGGCTTTT 1284  
61 ATGAGTGCAGAGATTAATAATCACTCAAGACATGCTGACGCCCAAGTGAAGCGGCTTTT 120  
1285 TCTGATGAAGAGAGAGTTCAGAGAGACTGCTGAGTGTGACAGTGAAGTGC 1344  
121 TCTGATGAAGAGAGAGTTCAGAGAGACTGCTGAGTGTGACAGTGAAGTGC 180  
1345 TCGACATTAGCCAGCCATACGTGTGTGCTGCGGAGTGTCTGAGTACAGAAAGCAGCG 1404  
181 TCGACATTAGCCAGCCATACGTGTGTGCTGCGGAGTGTCTGAGTACAGAAAGCAGCG 240  
1405 GCGAGAGCTCCGCACTGCGCCAGACCCGAGGCGGAGAGGCCCAAGGCCCTGAGG 1464  
241 GCGAGAGCTCCGCACTGCGCCAGACCCGAGGCGGAGAGGCCCAAGGCCCTGAGG 300  
1465 GATGACCCCTCCAGCTCCGTCAGCTGACGACAGAGTCCAGATTACGTGTGCTCTG 1524  
301 GATGACCCCTCCAGCTCCGTCAGCTGACGACGAGTCCAGATTACGTGTGCTCTG 357  
1525 CAGAGAGCCACGCCCTGTGTGACCTGCTTCCAGCCATGCCGAGCCGAGAGCGAG 1584  
358 CAGAGAGCCACGCCCTGTGTGACCTGCTTCCAGCCATGCCGAGCCGAGAGCGAG 417  
1585 GCGAGAGCGACCCGCGGTGTGCGCCCTCAGAGAGTGTGCTGCTGCTGCTGCTGCTG 1644  
418 GCGAGAGCGACCCGCGGTGTGCGCCCTCAGAGAGTGTGCTGCTGCTGCTGCTGCTG 477  
1645 CACCTGTACTGAGGCTGACCCGAGACCGGCTGCTACAGGCTGCTGCTGCTGCTGCTG 1704  
478 CACCTGTACTGAGGCTGACCCGAGACCGGCTGCTACAGGCTGCTGCTGCTGCTGCTG 537  
1705 CTCAACTGTGTGACCAAGTGTGTGAGAGGCTGCTGCAACACACAGTACAGTCAAGC 1764  
538 CTCAACTGTGTGACCAAGTGTGTGAGAGGCTGCTGCAACACACAGTACAGTCAAGC 597  
1765 ATCCGAAAGATTTCCTGCGACCAAGAGGTTTGAATGAAATCATGTTGACGAGAGC 1824  
598 ATCCGAAAGATTTCCTGCGACCAAGAGGTTTGAATGAAATCATGTTGACGAGAGC 657  
1825 CTC 1827  
658 CTC 660

## RESULT 38

AM957709 650 bp mRNA linear EST 01-JUN-2000  
LOCUS AM957709  
DEFINITION EST369779 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM957709

```

VERSION      AM957109.1  GI:8147392
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS      Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C.,
             Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and
             Quackenbush, J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
             metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
             The Institute for Genomic Research
             9712 Medical Center Dr., Rockville, MD 20850, USA
             Tel: 301 838 3528
             Fax: 301 838 0208
             Email: johnq@tigr.org
             Plate: 110
FEATURES     Seq primer: Reverse.
             Location/Qualifiers
             source          1..650
                           /organism="Homo sapiens"
                           /mol_type="mRNA"
                           /db_xref="taxon:9606"
                           /clone_lib="MAGE resequences, MAGE"
                           /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      15.6%; Score 418; DB 2; Length 650;
Best Local Similarity 100.0%; Pred. No. 5.3e-197;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2242 TTTACCTCTGAGTGAAGACCTTCCCAAGCCCCGGGGCCGAGCCCTCTCTGTTG 2301
DB      158 TTTACCTCTGAGTGAAGACCTTCCCAAGCCCCGGGGCCGAGCCCTCTCTGTTG 217
QY      2302 AGCGCTGGGCGAGGCTGTGTGGCATGACGACAGAGCAAGACCTTCTGTAAATAC 2361
DB      218 AGCGCTGGGCGAGGCTGTGTGGCATGACGACAGAGCAAGACCTTCTGTAAATAC 277
QY      2362 GGCCTGCTCCGCGAGAGGGGAGTTTGTCTTTTGTATCATTTTCCGAAATACAGTTAA 2421
DB      278 GGCCTGCTCCGCGAGAGGGGAGTTTGTCTTTTGTATCATTTTCCGAAATACAGTTAA 337
QY      2422 AGCAGAGTCTGTTTTCAGGAAAAGTTTCAAGGAGAGAGGGCAGTTTATCAAAAACATT 2481
DB      338 AGCAGAGTCTGTTTTCAGGAAAAGTTTCAAGGAGAGAGGGCAGTTTATCAAAAACATT 397
QY      2482 GTTTCAGAGAGAGGAGCATTAAGTTTACAGCCTTACAGAGAGTACCAATATCTGCTGCT 2541
DB      398 GTTTCAGAGAGAGGAGCATTAAGTTTACAGCCTTACAGAGAGTACCAATATCTGCTGCT 457
QY      2542 GGGAAAACCAAGCATTTTATATTTTATTTTATTTAGTTTGGTGTATCTCTTAA 2601
DB      458 GGGAAAACCAAGCATTTTATATTTTATTTTATTTAGTTTGGTGTATCTCTTAA 517
QY      2602 TAGAGTTTAAATGTCAAAACTGTAGCAAAATATATATTTATTTATTTCAAAATTG 2659
DB      518 TAGAGTTTAAATGTCAAAACTGTAGCAAAATATATATTTATTTATTTCAAAATTG 575

```

## RESULT 39

```

LOCUS      CN428708          414 bp      mRNA      linear      EST 16-MAY-2004
DEFINITION 17000532786421 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION  CN428708
VERSION    CN428708.1  GI:47416302
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE    Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS      1 (bases 1 to 414)
             Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Flak, G.J.,
             Li, Y., Xu, C., Pang, R., Quegler, K., Rao, M.S., Mandalam, R.,
             Lebkowski, J. and Stanton, L.W.
TITLE        Transcriptional characterization elucidates signaling networks that
             control human ES cell growth and differentiation
JOURNAL      Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT      Contact: Brandenberger R
             Regenerative Medicine
             Genon Corporation
             230 Constitution Drive, Menlo Park, CA 94025, USA
             Tel: 650 473 8658
             Fax: 650 473 7760
             Email: rbrandenberger@genon.com
             Insert Length: 414 Std Error: 0.00.
FEATURES     Location/Qualifiers
             source          1..414
                           /organism="Homo sapiens"
                           /mol_type="mRNA"
                           /db_xref="taxon:9606"
                           /tissue_type="embryonic stem cells, cell lines H1, H7, and
                           H9"
                           /clone_lib="GRN ES"
                           /note="oligo dt primed, full-length enriched cDNA library
                           from undifferentiated hES cell lines H1 (p32), H7 (p29),
                           and H9 (p26) maintained in feeder-free conditions"
ORIGIN
Query Match      15.5%; Score 414; DB 7; Length 414;
Best Local Similarity 100.0%; Pred. No. 5.3e-195;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      54 GTTCCGGGTTGGGCGCGGGGCGGGATGTGAATCCGATGAGCGCCCGAGAGAGGCAA 113
DB      1 GTTCCGGGTTGGGCGCGGGGCGGGATGTGAATCCGATGAGCGCCCGAGAGAGGCAA 60
QY      114 GCAAGTCGCGCGCGCGAGCCCTGGGGGAGCGGCTCCGCTGCGGGGCGGAGAGAGGCGCA 173
DB      61 GCAAGTCGCGCGCGCGAGCCCTGGGGGAGCGGCTCCGCTGCGGGGCGGAGAGAGGCGCA 120
QY      174 GCGGACGTCCTCTGAGGAGAGCGGAGTGAACATCGGCGGAGACAGAGTTGCGACCT 233
DB      121 GCGGACGTCCTCTGAGGAGAGCGGAGTGAACATCGGCGGAGACAGAGTTGCGACCT 180
QY      234 TTCTTCTCCAGCAATTAATCTGCTCTGGAATCACTGTAAATTTGATGATGAAA 293
DB      181 TTCTTCTCCAGCAATTAATCTGCTCTGGAATCACTGTAAATTTGATGATGAAA 240
QY      294 ATCAGGTGAGTGAACATCGGAGATACAGACACAGTGAACAGGATTAACAAGCTGAA 353
DB      241 ATCAGGTGAGTGAACATCGGAGATACAGACACAGTGAACAGGATTAACAAGCTGAA 300
QY      354 GGTGTTTAAAGAGCAGACATGCTTTTACAGCTGGGAGTGTATCTATCTGTTGATGAG 413
DB      301 GGTGTTTAAAGAGCAGACATGCTTTTACAGCTGGGAGTGTATCTATCTGTTGATGAG 360
QY      414 GAAAGATGAACCGGAACACACAGCTGGCATCTCTATGAACTTTTAAGTGAAGA 467
DB      361 GAAAGATGAACCGGAACACACAGCTGGCATCTCTATGAACTTTTAAGTGAAGA 414

```

## RESULT 40

```

LOCUS      CN428710          614 bp      mRNA      linear      EST 16-MAY-2004
DEFINITION 17000455366473 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION  CN428710
VERSION    CN428710.1  GI:47416304
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
1 (baes 1 to 614)	Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandlam, R., Lebkowski, J. and Stanton, L.W.	Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation	Nat. Biotechnol. 22 (6), 707-716 (2004)	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 614 Std Error: 0.00.	Location/Qualifiers 1..614 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /tissue_type="embryonic stem cells, cell lines H1, H7, and H9" /note="Toligo dT-primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p23), and H9 (p26) maintained in feeder-free conditions"
ORIGIN	Query Match	Best Local Similarity	Matches	15.4%; Score 412; DB 7; Length 614;	100.0%; Pred. No. 5,3e-194; Mismatches 0; Indels 0; Gaps 0;
QY	83 GAATCCGATGAGCGGCCCGAGAGGCAAGCAAGTCCGCCGCCGACCTCGGGAGC	142			
Db	39 GAATCCCGATGAGCGGCCCGAGAGGCAAGCAAGTCCGCCGCCGACCTCGGGAGC	98			
QY	143 GGCTCTGCGTCTGGGCGCGAGAGGGCGAGCGCGACGTCCTCTGAGGAAGCGGGAGT	202			
Db	99 GGCTCTGCGTCTGGGCGCGAGAGGGCGAGCGCGACGTCCTCTGAGGAAGCGGGAGT	158			
QY	203 GGACCATCGGCGGAGAGCGAGGTTCCGACCTTTCCTCCCGACAACTAACTGTCTGTG	262			
Db	159 GGACCATCGGCGGAGAGCGAGGTTCCGACCTTTCCTCCCGACAACTAACTGTCTGTG	218			
QY	263 GAGATCACTGTGAATTTGTAGTGTGATGAAAAATCGAGTCACTGGAAGATTAACA	322			
Db	219 GAGATCACTGTGAATTTGTAGTGTGATGAAAAATCGAGTCACTGGAAGATTAACA	278			
QY	323 GCACCAAGTGAACAGTGAATTAACAAGCTGAAGGTTGTTAAGAAGACAGATGCCCTTAC	382			
Db	279 GCACCAAGTGAACAGTGAATTAACAAGCTGAAGGTTGTTAAGAAGACAGATGCCCTTAC	338			
QY	383 AGACTGGGAGTGTCACTTACTTGTGTGTACAGGAAGATGAACCGGAACAACGTTGCAT	442			
Db	339 AGACTGGGAGTGTCACTTACTTGTGTGTGTACAGGAAGATGAACCGGAACAACGTTGCAT	398			
QY	443 ACCCTATGAATCTTAAAGTGAAGAAAGCAAGGATGACCAAGAACTCTTGA	494			
Db	399 ACCCTATGAATCTTAAAGTGAAGAAAGCAAGGATGACCAAGAACTCTTGA	450			
RESULT 41	LOCUS	635 bp	mRNA	linear	EST 21-MAR-2001
DEFINITION	602328653P1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650348 5',				
ACCESSION	CG481936				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 635)	NIH-MGC	http://mgc.nci.nih.gov/.		
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Straubeberg, Ph.D.			
	Email: gga@rs-remail.nih.gov			
	Tissue Procurement: ATCC			
	cDNA Library Preparation: Ling Hong/Rubin Laboratory			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)			
	Clone Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNLN at:			
	http://Image.lnl.gov			
	Plate: L16M431 row: j column: 13			
	High quality sequence stop: 635.			
	Location/Qualifiers			
	1. 635			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:4650348"			
	/tissue_type="choriocarcinoma"			
	/lab_host="DH10B (phage-resistant)"			
	/clone_id="N1H MGC 21"			
	/note="Organ: placenta; Vector: pOT87; Site_1: XhoI;			
	Site_2: EcoRI; cDNA made by oligo-dt priming.			
	Directionally cloned into EcoRI/XhoI sites using the			
	following 5' adaptor: GGCAAGAG(g). Size-selected >500bp			
	for average insert size 1.8kb. Library constructed by			
	Ling Hong in the laboratory of Gerald M. Rubin (University			
	of California, Berkeley) using ZAP-cDNA synthesis kit			
	(Stratagene) and Superscript II RT (Life Technologies)."			

  

ORIGIN	Query Match	15.4%; Score 412; DB 4; Length 635;
	Best Local Similarity 100.0%; Pred. No. 5.3e-194;	
	Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	83 GAATCCCGATGAGACGGGCCGAGAGGCAAGCATGTGGCGCCGACGCTTGGGAC	142
Db	2 GAATCCCGATGAGACGGGCCGAGAGGCAAGCATGTGGCGCCGACGCTTGGGAC	61
OY	143 GGCTCTCGCGCTGTGGGCGGAGGAGCGGACGCGACGTCCTCTGAGAAAGCGGAGT	202
Db	62 GGCTCTCGCGCTGTGGGCGGAGGAGCGGACGCGACGTCCTCTGAGAAAGCGGAGT	121
OY	203 GGACCATCGGCGGAGACGAGGTTGCCACTTTCTTCCCAAGATAACTGCTCTG	262
Db	122 GGACCATCGGCGGAGACGAGGTTGCCACTTTCTTCCCAAGATAACTGCTCTG	181
OY	263 GAGATCACTGTAGAAATTGTAGTGTGATGAAAAATCAGGTCAGTGATCACTGAAATACCA	322
Db	182 GAGATCACTGTAGAAATTGTAGTGTGATGAAAAATCAGTCACTGAAATACCA	241
OY	323 GCACCAAGTGAACAGTATTAACAAGTGAAGTTGTTAAGAAACAGACATGCTTTTAC	382
Db	242 GCACCAAGTGAACAGTATTAACAAGTGAAGTTGTTAAGAAACAGACATGCTTTTAC	301
OY	383 AGACTGGGAGATGTATCTACTCTGTGTACAGAAAGAAATGAACCGGAAACAAAGCGGAT	442
Db	302 AGACTGGGAGATGTATCTACTCTGTGTACAGAAAGAAATGAACCGGAAACAAAGCGGAT	361
OY	443 ACCTCTATGATCTTTAAGTGAAGAAGGATGATGACAAAGATCTTTGA	494
Db	362 ACCTCTATGATCTTTAAGTGAAGAAGGATGATGACAAAGATCTTTGA	413

  

RESULT 42	BF109790/c	490 bp	mRNA	linear	EST 20-OCT-2000
LOCUS	BF109790	7169804.x1	Soares_NSF_P8_9M_OT_PA_P_S1	Homo sapiens	cDNA clone
DEFINITION	IMAGE:3526543 3', mRNA sequence.				
ACCESSION	BF109790				



VERSION BF109790.1 GI:10939480  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 490)  
 NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 TITLE Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 This clone is available royalty-free through LNLN; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Seq primer: -40UP from Glibco  
 High quality sequence stop: 460.  
 Location/Qualifiers  
 1..490  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3526543"  
 /lab\_host="DH10B"  
 /note="Organ: Soares NSF F8 9W OT PA P S1"  
 /note="Organ: pooled; Vector: pFRTD-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 Equal amounts of plasmid DNA from five normalized libraries were mixed, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NBHF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HF-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHF pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and W. Fatima Bonaldo."

## ORIGIN

Query Match 15.3%; Score 409; DB 2; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-192; Indels 0; Gaps 0;  
 Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2251 TGAAGTGAACCTCCAGAGCCCGGAGCCCGCTCTGCTGAGCGCTGCG 2310  
 409 TGAAGTGAACCTCCAGAGCCCGGAGCCCGCTCTGCTGAGCGCTGCG 350  
 2311 CAGGGCTCGTGTGATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2370  
 349 CAGGGCTCGTGTGATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 290  
 2371 GCGGAG 2430  
 289 GCGGAG 230  
 2431 CTGTTTTCAG 2490  
 229 CTGTTTTCAG 170  
 2491 GAAAG 2550  
 169 GAAAG 110  
 2551 ACAG 2610  
 109 ACAG 50  
 2611 AATGTCAAACTGTAGACAAATATATATATATATATATATATATATATATG 2659

Db 49 AATGTCAAACTGTAGACAAATATATATATATATATATATATATATATG 1

RESULT 43  
 LOCUS BU153737  
 DEFINITION BU153737  
 AGENCOUNT 7889369 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:6142354  
 5', mRNA sequence.  
 ACCESSION BU153737  
 VERSION BU153737.1 GI:22667269  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 881)  
 NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
 Plate: L1AM13464 row: a column: 11  
 High quality sequence stop: 621.  
 Location/Qualifiers  
 1..881  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6142354"  
 /tissue\_type="retinoblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

## FEATURES

source

## ORIGIN

Query Match 15.1%; Score 404; DB 5; Length 881;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-190; Indels 0; Gaps 0;  
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 30 TTGACAGCG 89  
 1 TTGACAGCG 60  
 90 GATGAGCT 149  
 61 GATGAGCT 120  
 150 GCGTTCG 209  
 121 GCGTTCG 180  
 210 CCGGCGGAG 269  
 181 CCGGCGGAG 240  
 270 CTGTGAATTTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 329  
 241 CTGTGAATTTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 330 TGGAAAGTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 389  
 301 TGGAAAGTGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 390 GGATGTCATCTACTTGGTGTACAGAGAAATGAACCGAACA 433  
 DB 361 GAGTTCATCTACTTGGTGTACAGAGAAATGAACCGAACA 404

## RESULT 44

LOCUS BX409790 940 bp mRNA linear EST 01-MAY-2004  
 DEFINITION BX409790 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
 CS0DP017Y11.9 5-PRIME, mRNA sequence.

ACCESSION BX409790  
 VERSION BX409790.2 GI:46933826  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 940)

AUTHORS Li W.B., Gruber C., Jessee J., and Polayes D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)

COMMENT On May 13, 2003 this sequence version replaced gi:30656944.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6792.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?c=CS0BAF0142C07\_AFO1300\_1&c=6792.r

## FEATURES

source 1. 940 Location/Qualifiers

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF017Y11.9"  
 /tissue\_type="FETAL BRAIN"  
 /dev\_stage="fetal"  
 /clone.lib="Homo sapiens FETAL BRAIN"  
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Query Match 14.7%; Score 394; DB 5; Length 940;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-185;  
 Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1772 AGAATTAAGTCTGCAACAGAGTTTGACATGAAAAACATGTTGACCGAGACCTTCGCG 1831  
 DB 199 AAGATTAAGTCTGCAACAGAGTTTGACATGAAAAACATGTTGACCGAGACCTTCGCG 258  
 QY 1832 CTCTCAGAGGGGAGAGTGTCTGCTGTGATTAAGAGATCAACGGAGACACCGTTCTGT 1891  
 DB 259 CTCTCAGAGGGGAGAGTGTCTGCTGTGATTAAGAGATCAACGGAGACACCGTTCTGT 318  
 QY 1892 GTTACTGCTGTGAGCTGCGCAGCTTCGTAAGCTGAACCTATCATGATCGGCAAAATTTC 1951  
 DB 319 GTTACTGCTGTGAGCTGCGCAGCTTCGTAAGCTGAACCTATCATGATCGGCAAAATTTC 378  
 QY 1952 CTGCTTCGAGAGTTGCGAGTGGCGGTAACTCCGCTCTGATCTGCTAAGTGGGCGGTAACT 2011  
 DB 379 CTGCTTCGAGAGTTGCGAGTGGCGGTAACTCCGCTCTGATCTGCTAAGTGGGCGGTAACT 438  
 QY 2012 GCGGCACTGAGGTAAGCTCAACGACCATGAAATTCATCATCTGTGAACAGACAA 2071

DB 439 GCGGCACTGAGGTAAGCTCAACGACCATGAAATTCATCATCTGTGAACAGACAA 498  
 QY 2072 GGTTCAAAACCTAAGCATCCAGAGGCTTGAAGCAGCTTTCAGCACTGAGAGTGAAGAG 2131  
 DB 499 GGTTCAAAACCTAAGCATCCAGAGGCTTGAAGCAGCTTTCAGCACTGAGAGTGAAGAG 558

QY 2132 CGTGTTCATTAATATACAGAGCAAGCAGCTCAAG 2165  
 DB 559 CGTGTTCATTAATATACAGAGCAAGCAGCTCAAG 592

## RESULT 45

LOCUS AI200393 475 bp mRNA linear EST 14-OCT-1998  
 DEFINITION qf6e04.x1 Soares\_fetal\_lung\_NbH19w Homo sapiens cDNA clone  
 IMAGE:1756926 3', mRNA sequence.

ACCESSION AI200393  
 VERSION AI200393.1 GI:3752999  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 475)

AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 CONTACT Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -400D from Gluco  
 High quality sequence stop: 474.

## FEATURES

source 1. 475 Location/Qualifiers

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1756926"  
 /dev\_stage="19 weeks"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone.lib="Soares fetal lung NbH19w"  
 /note="Organ: lung; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer  
 (5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTCTTTT-3'),  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M. Fatima Bonaldo. This library was constructed  
 from the same fetus as the fetal heart library, Soares  
 fetal heart NbH19w."

## ORIGIN

Query Match 14.4%; Score 386; DB 1; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-181;  
 Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2276 GGGGGCCGACCCGCTCTCTGTGAGCGCTGGGAGGCTTCGTGCGCATGACGACG 2335  
 DB 474 GGGGGCCGACCCGCTCTCTGTGAGCGCTGGGAGGCTTCGTGCGCATGACGACG 415  
 QY 2336 AGAGCAAGGCTTCTGTACATGCGGCGCTCCGCGAGAGGGGAGTTTGGCTCTT 2395  
 DB 414 AGAGCAAGGCTTCTGTACATGCGGCGCTCCGCGAGAGGGGAGTTTGGCTCTT 355  
 QY 2396 TGTACATTTTCGAAATCATAGTTAAAGCAGAGTCTGTTTCAGAAAAAGTTTCAAGGG 2455  
 DB 354 TGTACATTTTCGAAATCATAGTTAAAGCAGAGTCTGTTTCAGAAAAAGTTTCAAGGG 295

QY 2456 AGAAGGCAAGTTTCAAAAACATTGTTTCAGAGAGGAGCATTAAGTTTACAGCCTA 2515  
 DB 294 AGAAGGCAAGTTTATCAAAAACATTGTTTCAGAGAGGAGCATTAAGTTTACAGCCTA 235  
 QY 2516 CAGAGAGTACAAATATCTGCTGCTGGGAAAACACAGATTTTATCTATTTTATTT 2575  
 DB 234 CAGAGAGTACAAATATCTGCTGCTGGGAAAACACAGATTTTATCTATTTTATTT 175  
 QY 2576 TAATAGTTTGGTCTTATCTTCAATTAAGATTAAATGTCACAACTGTAGCACAAT 2635  
 DB 174 TAATAGTTTGGTCTTATCTTCAATTAAGATTAAATGTCACAACTGTAGCACAAT 115  
 QY 2636 ATATAATTATTAATTAACAATTGAC 2661  
 DB 114 ATATAATTATTAATTAACAATTGAC 89

RESULT 46  
 CN481126 618 bp mRNA linear EST 26-APR-2004  
 LOCUS hw03c10.y1 Human primary human ocular pericytes. Unamplified (hw)  
 DEFINITION Homo sapiens cDNA clone hw03c10.5', mRNA sequence.  
 ACCESSION CN481126 GI:46562630  
 VERSION CN481126  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Tsal, J.Y. and Wistow, G.  
 1 (bases 1 to 618)  
 Expressed sequence tag analysis of cultured primary human ocular  
 pericytes  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: gtaem@helix.nih.gov  
 Plate: 03 row: C column: 10  
 Seq primer: M13Rpl reverse primer (ABI).  
 Location/Qualifiers  
 1..618  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="hw03c10"  
 /cell\_type="pericytes"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human primary human ocular pericytes."  
 /note="Organ: Eye; Vector: pSPORT1; RNA was extracted from  
 primary human pericytes in culture. A directionally cloned  
 cDNA library in the pSPORT1 vector (Invitrogen) was  
 essentially following the protocols of the Superscript  
 Plasmid System full details of which are contained in the  
 manufacturer's instruction manual  
 (http://www.lifetechn.com/). First strand synthesis was  
 carried out using a Not I primer-adaptor  
 (5'-GACTAGTCTTAGATCGCAGCGGCCCTT)15-3'. cDNA was  
 cloned in Not I/Sal I sites. EST analysis was performed at  
 the NIH Intramural Sequencing Center (NISC)."

ORIGIN  
 Query Match 14.4%; Score 386; DB 7; Length 618;  
 Best Local Similarity 99.8%; Pred. No. 5.3e-181;  
 Matches 436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1806 AAACAGTTGACGAGAGCTGTGCTCTCCAGCGGAGAGTGTCTCTGTCTGATTA 1865

DB 115 AAACAGTTGACGAGAGCTGTGCTCTCCAGCGGAGAGTGTCTCTGTCTGATTA 174  
 QY 1866 CAGAGTACGAGAGAGACCGCTTCTGTGTATCTGTGTGCTGTGAGCTTCCGAGCT 1925  
 DB 175 CAGAGTACGAGAGAGACCGCTTCTGTGTATCTGTGTGCTGTGAGCTTCCGAGCT 234  
 QY 1926 GACCTATCAGTATCCGACAGAAATTCCTGCTCCGAGTTCAGAGTTCGCTTAATCCCG 1985  
 DB 235 GACCTATCAGTATCCGACAGAAATTCCTGCTCCGAGTTCAGAGTTCGCTTAATCCCG 294  
 QY 1986 TCTGACTGCTACTGAGGAGCGTAACTGCGCACTCAGGTGAAGCTCACACGCCATGA 2045  
 DB 295 TCTGACTGCTACTGAGGAGCGTAACTGCGCACTCAGGTGAAGCTCACACGCCATGA 354  
 QY 2046 ATTCAATCATATCTGTGAACAGACAGAGTTCAAAAATAAGCATTCAGAGGCTTCAGCA 2105  
 DB 355 ATTCAATCATATCTGTGAACAGACAGAGTTCAAAAATAAGCATTCAGAGGCTTCAGCA 414  
 QY 2106 GCTTTCAGACTGAGGTGAGAGGAGCGTGTATTAATTAAGAGACAGACAGCGTCAAG 2165  
 DB 415 GCTTTCAGACTGAGGTGAGAGGAGCGTGTATTAATTAAGAGACAGACAGCGTCAAG 474  
 QY 2166 GTGTTTTCACAGCCCTCTGAGGAGGAGGAGCGAGGCTTCGACAGGTGCTGTGGGTGA 2225  
 DB 475 GTGTTTTCACAGCCCTCTGAGGAGGAGGAGCGAGGCTTCGACAGGTGCTGTGGGTGA 534  
 QY 2226 CTCTTCTGTGAGCTTT 2242  
 DB 535 CTCTTCTGTGAGCTTT 551

RESULT 47  
 BG831289 824 bp mRNA linear EST 22-MAY-2001  
 LOCUS 602766120P1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:490816.5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG831289  
 VERSION BG831289.1 GI:14178876  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 824)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rcmail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: LNCM1814 row: G column: 05  
 High quality sequence stop: 745.  
 Location/Qualifiers  
 1..824  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:490816"  
 /tissue="IMR90:epithelial carcinoma cell line"  
 /lab\_host="MDH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_42"  
 /note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;  
 Site: 2: EcoRI; cDNA made by oligo-dT priming.  
 directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling

## ORIGIN

Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library. |"

Query Match 14.3%; Score 382; DB 4; Length 824;  
Best Local Similarity 99.8%; Pred. No. 5.3e-179;  
Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1810 ATGTGACCGAGAGCGCTGCTCCAGCGGGAGTGTTCGTGCTGATTAACGA 1869  
1810 ATGTGACCGAGAGCGCTGCTCCAGCGGGAGTGTTCGTGCTGATTAACGA 61  
2 ATGTGACCGAGAGCGCTGCTCCAGCGGGAGTGTTCGTGCTGATTAACGA 61  
1870 GTACGCGGAGACCGCTTCTGTGTTACTGCTGTGGCTTCCGAGCTTCCGTAGTGAAC 1929  
62 GTACGCGGAGACCGCTTCTGTGTTACTGCTGTGGCTTCCGAGCTTCCGTAGTGAAC 121

1930 TATCATATCGGACAAACATTCCTGCTCCAGTTCGAGTTCGAGTTCGATTCCTCT 1989  
122 TATCATATCGGACAAACATTCCTGCTCCAGTTCGAGTTCGAGTTCGATTCCTCT 181

1990 GACTGCTACTGGGGCCGTAACTGCGCACTCAGTGAAGCTCACGAGCCATGAATTC 2049  
182 GACTGCTACTGGGGCCGTAACTGCGCACTCAGTGAAGCTCACGAGCCATGAATTC 241

2050 AATCATATCTGTGAACAGCAAGTTCAAAAATAAGCATCCAGAGCCCTGAGAGCTT 2109  
242 AATCATATCTGTGAACAGCAAGTTCAAAAATAAGCATCCAGAGCCCTGAGAGCTT 301

2110 TCAGCACTGAGAGTGAAGAGCGCTTTTAAATACAGACAAACGCTCAAGGTGT 2169  
302 TCAGCACTGAGAGTGAAGAGCGCTTTTAAATACAGACAAACGCTCAAGGTGT 361

2170 TTTCACAGCCCTGAGGAGGAGGAGCGAGGCTCTCCGACAGTGTCTGGGGTACTCT 2229  
362 TTTCACAGCCCTGAGGAGGAGGAGCGAGGCTCTCCGACAGTGTCTGGGGTACTCT 421

2230 TCTGTGAGCTTT 2242  
422 TCTGTGAGCTTT 434

RESULT 48  
BI460002 797 bp mRNA linear EST 21-AUG-2001  
LOCUS 603201532F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5267575 5',  
DEFINITION mRNA sequence.  
ACCESSION BI460002  
VERSION BI460002.1 GI:15250658  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
JOURNAL NIH-MGC http://mgi.nci.nih.gov/  
COMMENT Unpublished (1999)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1674 row: P column: 08  
High quality sequence stop: 793.  
Location/Qualifiers  
1. 797

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5267575"  
/lab\_host="DH10B"  
/clone\_id="NIH\_MGC\_97"  
/note="Organ: testis; Vector: pBluescript (modified)  
pBluescript KS+; Site 1: BamHI; Site 2: SalI-XhoI  
(gtcgag); Oligo-dT primed using primer  
5'-TTTTTTTTTTTTT-3', size-selected for average  
insert size 2.2 kb and normalized to R0.5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: this is  
a NIH\_MGC Library."

Query Match 14.2%; Score 381; DB 4; Length 797;  
Best Local Similarity 100.0%; Pred. No. 1.7e-178;  
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

53 GATTCCGGGTCGAGCGGGGAGTGAATCCGATGAGCGCCGAGAGAGCA 112  
9 GATTCCGGGTCGAGCGGGGAGTGAATCCGATGAGCGCCGAGAGAGCA 68

113 AGCACTGCGCGCGCGAGCCCTGAGGAGCGCTCTGCGCGCGAGAGAGCGG 172  
69 AGCACTGCGCGCGCGAGCCCTGAGGAGCGCTCTGCGCGCGAGAGAGCGG 128

173 AGCGGAGCTCTCTGAGAGAGCGGAGTGAACATCGCGCGAGAGAGAGCGG 232  
129 AGCGGAGCTCTCTGAGAGAGCGGAGTGAACATCGCGCGAGAGAGAGCGG 188

223 TTTCCTCCCGAGCAATTAATCTGCTCTGAGAGTCACTGAGATTGATGATGAAA 292  
189 TTTCCTCCCGAGCAATTAATCTGCTCTGAGAGTCACTGAGATTGATGATGAAA 248

293 AATCAGTCACTGAGCACTGAGAGTACAGACACAGTGAACAGTGAACAGCTGA 352  
249 AATCAGTCACTGAGCACTGAGAGTACAGACACAGTGAACAGTGAACAGCTGA 308

353 AGTTGTTAAGAGCAAGATGCTTTCACAGTCTGGGGATGATCTGATGATGATCA 412  
309 AGTTGTTAAGAGCAAGATGCTTTCACAGTCTGGGGATGATCTGATGATGATCA 368

413 GGAAGATGAACCGGACACA 433  
369 GGAAGATGAACCGGACACA 389

RESULT 49  
BM602324 1070 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT\_6460403 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:555833  
DEFINITION 5', mRNA sequence.  
ACCESSION BM602324  
VERSION BM602324.1 GI:19119147  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
JOURNAL NIH-MGC http://mgi.nci.nih.gov/  
COMMENT Unpublished (1999)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation

FEATURES  
source























[illegible]

## RESULT 1

; Sequence 18, Application US/07847010

**; Patent No. 5693495**

**GENERAL INFORMATION:**

APPLICANT: Breiteneder, Heimo

APPLICANT: Kelkerstorfer, Arnold  
1001 1st St. S. #315  
Minneapolis, MN 55415

APPLICANT: Valencia, Rudolf

APPLICANT: Hoffmann - Sommergrün

APPLICANT: Kraft, Dietrich

APPLICANT: Rumbold, Helmut

APPLICANT: Scheiner, Otto

APPLICANT: Ebner, Christof

APPLICANT: Ferreira, Fatima

**TITLE OF INVENTION: Allergens of**

TITLE OF INVENTION: Applications

```
NUMBER OF SEQUENCES: 23
```

CORRESPONDENCE ADDRESS: Dr. J. C. Edwards  
; ADDRESS: Dr. J. C. Edwards

ADDRESSEE: FEMITE & ELLIOTTS  
STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.
CURRENT APPLICATION DATA:

```

CORRECTION DATA: APPLICATION NUMBER: IIS/07/847

RECEIVED DATE: 01-JUN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jones III, Harry C

REGISTRATION NUMBER: 20,280

REFERENCE/DOCKET NUMBER: 6530-;

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE  
TELEPHONE FOR STD NO. 19.

; INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:

SEQUENCE CIPHERS:

```

; LENGTH: 860 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: hazel (Corylus sp.)
; IMMEDIATE SOURCE:
; LIBRARY: POLLEN FROM ALLERGEN AB, ENGELHOLM, SWEDEN
US-07-847-010-18

Query Match          1.0%; Score 26; DB 1; Length 860;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches    26; Conservative   0; Mismatches      0; Indels      0; Gaps      0

QY       2654 AAATTGACAAAAAAAAAAAAAAAAAAAAA 2679
           |||||
DB        754 AAATTGACAAAAAAAAAAAAAAAAAAAAA 729

```

; LENGTH: 258  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-11988

Query Match 0.9%; Score 24; DB 4; Length 258;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679  
DB 220 ATTGACAAAAA 243

RESULT 4  
5198347-3  
; Patent No. 5198347  
; APPLICANT: Miller, Louis H.; Adams, John H.; Kaslow,  
; Davis C.; Fang, Xiangdou  
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND  
; PLASMODIUM KNOWLES DUFFY RECEPTOR  
; NUMBER OF SEQUENCES: 27  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/554,837  
; FILING DATE: 20-JUL-1990  
; SEQ ID NO: 3  
; LENGTH: 3157  
5198347-3

Query Match 0.9%; Score 24; DB 6; Length 3157;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679  
DB 3088 ATTGACAAAAA 3111

RESULT 5  
US-09-621-976-17572  
; Sequence 17572, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 17572  
; LENGTH: 443  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-17572

Query Match 0.9%; Score 23; DB 4; Length 443;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679  
DB 420 TTGACAAAAA 442

RESULT 6  
US-09-614-912-161  
; Sequence 161, Application US/09614912  
; Patent No. 6677502  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve

; APPLICANT: Rafalski, Antoni  
; APPLICANT: Orozco, Buddy  
; APPLICANT: Miao, Gou-Hau  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian Ming  
; APPLICANT: Sakai, Hajime  
; APPLICANT: Weng, Zude  
; APPLICANT: Caimi, Perry G  
; APPLICANT: Anderson, Shawn  
; TITLE OF INVENTION: Plant Metabolism Genes  
; FILE REFERENCE: B01378 US NA  
; CURRENT APPLICATION NUMBER: US/09/614,912  
; CURRENT FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: 60/143,401  
; PRIOR FILING DATE: 1999-07-12  
; PRIOR APPLICATION NUMBER: 60/143,412  
; PRIOR FILING DATE: 1999-07-12  
; PRIOR APPLICATION NUMBER: 60/146,650  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 60/170,906  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: 60/172,959  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/172,946  
; PRIOR FILING DATE: 1999-12-21  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 161  
; LENGTH: 1185  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-614-912-161

Query Match 0.9%; Score 23; DB 4; Length 1185;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAA 2679  
DB 1163 TTGACAAAAA 1185

RESULT 7  
US-09-412-102-1  
; Sequence 1, Application US/09412102  
; Patent No. 6228992  
; GENERAL INFORMATION:  
; APPLICANT: JESSEN, HOLLY J  
; APPLICANT: MEYER, TERRY B  
; TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF  
; TITLE OF INVENTION: NEMATODES IN PLANTS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: W. Murray Spruill (Alston & Bird, LLP)  
; STREET: 3605 Glenwood Ave. Suite 310  
; CITY: Raleigh  
; STATE: NC  
; COUNTRY: US  
; ZIP: 27622  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/412,102  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/217,787  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-18P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919 420 2202  
TELEFAX: 919 881 3175  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Zea mays  
IMMEDIATE SOURCE:  
CLONE: P12217  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 146..994  
US-09-412-102-1

Query Match 0.9%; Score 23; DB 3; Length 1347;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2679  
DB 1325 TTGACAAAAAAAAAAAAAAAAA 1347

RESULT 8  
US-09-217-787-1  
Sequence 1, Application US/09217787  
Patent No. 6284948  
GENERAL INFORMATION:  
APPLICANT: JESSEN, HOLLY J  
APPLICANT: MEYER, TERRY E  
TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF  
TITLE OF INVENTION: NEMATODES IN PLANTS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)  
STREET: 3605 Glenwood Ave. Suite 310  
CITY: Raleigh  
STATE: NC  
COUNTRY: US  
ZIP: 27622  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/217,787  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-18P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919 420 2202  
TELEFAX: 919 881 3175  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Zea mays  
IMMEDIATE SOURCE:

CLONE: P12217  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 146..994  
US-09-217-787-1

Query Match 0.9%; Score 23; DB 3; Length 1347;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2679  
DB 1325 TTGACAAAAAAAAAAAAAAAAA 1347

RESULT 9  
US-08-484-101B-35  
Sequence 35, Application US/08484101B  
Patent No. 5624868  
GENERAL INFORMATION:  
APPLICANT: California Institute of Technology  
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO  
TITLE OF INVENTION: ETHYLENE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richard F. Treccartin  
STREET: 3400 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,101B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/  
FILING DATE: 01-JUL-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/086,555  
FILING DATE: 01-JUL-1993  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Treccartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-57515-2/RPT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2405 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 288..2196  
US-08-484-101B-35

Query Match 0.9%; Score 23; DB 1; Length 2405;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2679  
DB 2378 TTGACAAAAAAAAAAAAAAAAA 2400

## RESULT 10

US-08-484-101B-49

; Sequence 49, Application US/08484101B  
; Patent No. 5824868

; GENERAL INFORMATION:

; APPLICANT: California Institute of Technology  
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Richard F. Treacartin  
; STREET: 3400 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,101B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 800

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/086,555

; FILING DATE: 01-JUL-1993

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Treacartin, Richard F.

; REGISTRATION NUMBER: 31,801

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; INFORMATION FOR SEQ ID NO: 49:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2405 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 288..2196

; US-08-484-101B-49

Query Match 0.9%; Score 23; DB 1; Length 2405;  
Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACCAAAAAAAAAAAAAAAAAA 2679

DB 2378 TTGACCAAAAAAAAAAAAAAAAAA 2400

## RESULT 11

US-08-714-524D-35

; Sequence 35, Application US/08714524D  
; Patent No. 6294716

; GENERAL INFORMATION:

; APPLICANT: Meyerowitz, Elliott M

; APPLICANT: Chang, Caren

; APPLICANT: Bleecker, Anthony B

; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE

; FILE REFERENCE: a-57515-4

; CURRENT APPLICATION NUMBER: US/08/714,524D

; CURRENT FILING DATE: 1996-09-16  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO: 35  
; LENGTH: 2405  
; TYPE: DNA  
; ORGANISM: Lycopersicon esculentum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (288)..(2195)  
; US-08-714-524D-35Query Match 0.9%; Score 23; DB 3; Length 2405;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACCAAAAAAAAAAAAAAAAAA 2679

DB 2378 TTGACCAAAAAAAAAAAAAAAAAA 2400

## RESULT 12

US-08-714-524D-49

; Sequence 49, Application US/08714524D  
; Patent No. 6294716

; GENERAL INFORMATION:

; APPLICANT: Meyerowitz, Elliott M

; APPLICANT: Chang, Caren

; APPLICANT: Bleecker, Anthony B

; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE

; FILE REFERENCE: a-57515-4

; CURRENT APPLICATION NUMBER: US/08/714,524D

; CURRENT FILING DATE: 1996-09-16

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Patent Ver. 2.1

; SEQ ID NO: 49

; LENGTH: 2405

; TYPE: DNA

; ORGANISM: Lycopersicon esculentum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (288)..(2195)  
US-08-714-524D-49Query Match 0.9%; Score 23; DB 3; Length 2405;  
Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACCAAAAAAAAAAAAAAAAAA 2679

DB 2378 TTGACCAAAAAAAAAAAAAAAAAA 2400

RESULT 13  
US-09-270-767-8213  
; Sequence 8213, Application US/09270767  
; Patent No. 6703451

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: Patent Ver. 2.0

; SEQ ID NO: 8213

; LENGTH: 121

; TYPE: DNA

; ORGANISM: Drosophila melanogaster  
US-09-270-767-8213Query Match 0.8%; Score 22; DB 4; Length 121;  
Best Local Similarity 100.0%; Pred. No. 12;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 93 TGACAAAAAAAAAAAAAAAAAAAA 114

## RESULT 14

US-09-270-767-23495  
; Sequence 23495, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23495  
; LENGTH: 121  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-23495

Query Match 0.8%; Score 22; DB 4; Length 121;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 93 TGACAAAAAAAAAAAAAAAAAAAA 114

## RESULT 15

US-08-244-113-18  
; Sequence 18, Application US/08244113  
; Patent No. 5455181  
; GENERAL INFORMATION:  
; APPLICANT: Strube, Karl-Hermann  
; APPLICANT: Bialojan, Siegfried  
; APPLICANT: Kroegeer, Burkhard  
; APPLICANT: Friedrich, Thomas  
; TITLE OF INVENTION: No. 5455181el thrombin-inhibitory proteins from terrestrial  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Keil & Weinkauff  
; STREET: 1101 Connecticut Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; OPERATING SYSTEM: MS-DOS version 6.0  
; SOFTWARE: Wordperfect version 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/244,113  
; FILING DATE:

CLASSIFICATION: 530  
; CLASSIFICATION: C07K 73/10  
; CLASSIFICATION: A61K 37/64  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP92/02661  
; FILING DATE: 19-NOV-1992  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 277 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-244-113-18

Query Match 0.8%; Score 22; DB 1; Length 277;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 224 TGACAAAAAAAAAAAAAAAAAAAA 245

## RESULT 16

US-09-280-116-159/C  
; Sequence 159, Application US/09280116A  
; Patent No. 6331427  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
; FILE REFERENCE: 5800-24, 035800/176965  
; CURRENT APPLICATION NUMBER: US/09/280,116A  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 159  
; LENGTH: 424  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Lon family of ATP-dependent proteases  
US-09-280-116-159

Query Match 0.8%; Score 22; DB 3; Length 424;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 22 TGACAAAAAAAAAAAAAAAAAAAA 1

## RESULT 17

US-09-311-021-161  
; Sequence 161, Application US/09311021  
; Patent No. 6706869  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agoestino, Michael J.  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
; FILE REFERENCE: GI 6300-11A  
; CURRENT APPLICATION NUMBER: US/09/311,021  
; CURRENT FILING DATE: 1999-05-13  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 161  
; LENGTH: 622  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-311-021-161

Query Match 0.8%; Score 22; DB 4; Length 622;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 589 TGACAAAAAAAAAAAAAAAAAAAA 610

RESULT 18  
US-09-205-258-172



```
; Sequence 172, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,839
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
```

---

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 172
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-205-258-172

Query Match          0.8%; Score 22; DB 4; Length 786;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2658 TGCACAAAAAAAAAAAAAAAAA 2679
Db       760 TGCACAAAAAAAAAAAAAAAAA 781

RESULT 19
US-09-270-767-5937
; Sequence 5937, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5937
; LENGTH: 809
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5937

Query Match          0.8%; Score 22; DB 4; Length 809;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2658 TGCACAAAAAAAAAAAAAAAAA 2679
Db       787 TGCACAAAAAAAAAAAAAAAAA 808

RESULT 20
US-09-270-767-21219
; Sequence 21219, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 21219
; LENGTH: 809
```

```

; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; US-09-270-767-21219

Query Match      0.8%; Score 22; DB 4; Length 809;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2658 TGCACAAAAAAAAAAAAAAAAA 2679
DB      787 TGCACAAAAAAAAAAAAAAAAA 808

RESULT 21
US-09-205-258-190/c
; Sequence 190, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06

```

```

; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 190
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (144)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-205-258-190

Query Match      0.8%; Score 22; DB 4; Length 906;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2657 TTGACAAAAAAAAAAAAAAAAA 2678
DB      843 TTGACAAAAAAAAAAAAAAAAA 822

RESULT 22
US-09-475-316A-24
; Sequence 24, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINO/LARICIPRESTINOL REDUCTASES.
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE
; FILE REFERENCE: WSUR-1-13793
; CURRENT FILING DATE: 1999-12-30
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653

```

PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: PCT/US97/20391  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/054,380  
PRIOR FILING DATE: 1997-07-31  
PRIOR APPLICATION NUMBER: 60/030,522  
PRIOR FILING DATE: 1996-11-08  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 914  
TYPE: DNA  
ORGANISM: *Thuja plicata*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (94)..(669)  
US-09-475-316A-24

Query Match 0.8%; Score 22; DB 3; Length 914;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2678  
|||||  
DB 893 TTGACAAAAAAAAAAAAAAAAA 914

RESULT 23  
US-09-704-640-24  
Sequence 24, Application US/09704640  
Patent No. 6635459  
GENERAL INFORMATION:  
APPLICANT: Lewis, No. 6635459man G.  
APPLICANT: Davin, Laurence B.  
APPLICANT: Dinkova-Kostova, Albena T.  
APPLICANT: Fujita, Masayuki  
APPLICANT: Gang, David R.  
APPLICANT: Sarkanen, Stimo  
APPLICANT: Ford, Joshua D  
TITLE OF INVENTION: RECOMBINANT PINGRESINOL/LARICRESINOL REDUCTASE,  
FILE REFERENCE: WSUR-1-16492  
CURRENT APPLICATION NUMBER: US/09/704,640  
CURRENT FILING DATE: 2000-11-02  
PRIOR APPLICATION NUMBER: 09/475,316  
PRIOR FILING DATE: 1999-12-30  
PRIOR APPLICATION NUMBER: 09/307,653  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: PCT/US97/20391  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/054,380  
PRIOR FILING DATE: 1997-07-31  
PRIOR APPLICATION NUMBER: 60/030,522  
PRIOR FILING DATE: 1996-11-08  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 914  
TYPE: DNA  
ORGANISM: *Thuja plicata*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (94)..(669)  
US-09-704-640-24

Query Match 0.8%; Score 22; DB 4; Length 914;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2678  
|||||  
DB 893 TTGACAAAAAAAAAAAAAAAAA 914

RESULT 24  
US-09-183-861-19  
Sequence 19, Application US/09183861  
Patent No. 6365165  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Skeiky, Yasir A.W.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/183,861  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/022,765  
FILING DATE: 12-FEB-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.420C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: *Leishmania major*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 71..523  
US-09-183-861-19

Query Match 0.8%; Score 22; DB 3; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 991 TGACAAAAAAAAAAAAAAAAA 1012

RESULT 25  
US-09-022-765-19  
Sequence 19, Application US/09022765  
Patent No. 6375955  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Skeiky, Yasir A.W.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE THERAPY AND

NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,765  
FILING DATE: 12-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.420C3  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1019 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Leishmania major  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 71..523  
US-09-022-765-19

Query Match 0.8%; Score 22; DB 3; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 991 TGACAAAAAAAAAAAAAAAAAAAA 1012

RESULT 26  
US-09-551-974A-19  
Sequence 19, Application US/09551974A  
Patent No. 6500437  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, David C.  
APPLICANT: Skeiky, Yasir A.W.  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
FILE REFERENCE: 210121.420C5  
CURRENT APPLICATION NUMBER: US/09/551,974A  
CURRENT FILING DATE: 2000-04-14  
NUMBER OF SEQ ID NOS: 101  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 1019  
TYPE: DNA  
ORGANISM: Leishmania major  
US-09-551-974A-19

Query Match 0.8%; Score 22; DB 4; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 991 TGACAAAAAAAAAAAAAAAAAAAA 1012

RESULT 27  
US-09-565-501A-19  
Sequence 19, Application US/09565501A  
Patent No. 6607731  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, David C.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Bhatia, Ajay  
APPLICANT: Coler, Rhea  
APPLICANT: Peter Probst  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
FILE REFERENCE: 210121.420C6  
CURRENT APPLICATION NUMBER: US/09/565,501A  
CURRENT FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 1019  
TYPE: DNA  
ORGANISM: Leishmania major  
US-09-565-501A-19

Query Match 0.8%; Score 22; DB 4; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||  
DB 991 TGACAAAAAAAAAAAAAAAAAAAA 1012

RESULT 28  
US-09-639-206A-19  
Sequence 19, Application US/09639206A  
Patent No. 6613337  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, David C.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Bhatia, Ajay  
APPLICANT: Coler, Rhea  
APPLICANT: Probst, Peter  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
FILE REFERENCE: 210121.420C7  
CURRENT APPLICATION NUMBER: US/09/639,206A  
CURRENT FILING DATE: 2000-08-14  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 1019  
TYPE: DNA  
ORGANISM: Leishmania major  
US-09-639-206A-19

Query Match 0.8%; Score 22; DB 4; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
|||||

Db 991 TGACAAAAA 1012

RESULT 29  
US-09-874-923-19  
Sequence 19, Application US/09874923  
Patent No. 6638517  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Campos-Neto, Antonio  
APPLICANT: Webb, John R.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Skeiky, Yasser A.W.  
APPLICANT: Bhatia, Ajay  
APPLICANT: Coler, Rhea  
APPLICANT: Probst, Peter  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
FILE REFERENCE: 210121.420C8  
CURRENT APPLICATION NUMBER: US/09/874,923  
CURRENT FILING DATE: 2001-06-04  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 1019  
TYPE: DNA  
ORGANISM: Leishmania major  
US-09-874-923-19

Query Match 0.8%; Score 22; DB 4; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAA 2679  
Db 991 TGACAAAAA 1012

RESULT 30  
US-08-798-841-19  
Sequence 19, Application US/08798841  
Patent No. 6709661  
GENERAL INFORMATION:  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/798,841  
FILING DATE: 12-FEB-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.420C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1019 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE: CDNA  
ORGANISM: Leishmania major  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 71..523  
US-08-798-841-19

Query Match 0.8%; Score 22; DB 4; Length 1019;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAA 2679  
Db 991 TGACAAAAA 1012

RESULT 31  
US-09-232-200-50  
Sequence 50, Application US/09232200A  
Patent No. 6288213  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Gimeno, Ruth E.  
TITLE OF INVENTION: FATY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WH197-21p3MB  
CURRENT APPLICATION NUMBER: US/09/232,200A  
CURRENT FILING DATE: 1999-01-14  
EARLIER APPLICATION NUMBER: 60/071,374  
EARLIER FILING DATE: 1998-01-15  
EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 50  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-232-200-50

Query Match 0.8%; Score 22; DB 3; Length 1173;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAA 2679  
Db 1140 TGACAAAAA 1161

RESULT 32  
US-09-232-197-50  
Sequence 50, Application US/09232197A  
Patent No. 6300096  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Gimeno, Ruth E.  
TITLE OF INVENTION: FATY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WH197-21p3MA  
CURRENT APPLICATION NUMBER: US/09/232,197A  
CURRENT FILING DATE: 1999-01-14  
EARLIER APPLICATION NUMBER: 60/071,374

EARLIER FILING DATE: 1998-01-15  
EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 50  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-232-195-50

Query Match 0.8%; Score 22; DB 3; Length 1173;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 1140 TGACAAAAAAAAAAAAAAAAAAAA 1161

## RESULT 33

US-09-232-201-50  
Sequence 50, Application US/09232201A  
Patent No. 6348321  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Gimeno, Ruth E.  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WHI97-21P3MC  
CURRENT APPLICATION NUMBER: US/09/232,201A  
EARLIER FILING DATE: 1999-01-14  
EARLIER APPLICATION NUMBER: 60/071,374  
EARLIER FILING DATE: 1998-01-15  
EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 50  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-232-201-50

Query Match 0.8%; Score 22; DB 3; Length 1173;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 1140 TGACAAAAAAAAAAAAAAAAAAAA 1161

## RESULT 34

US-09-232-195-50  
Sequence 50, Application US/09232195A  
Patent No. 6657049  
GENERAL INFORMATION:  
APPLICANT: Stahl, Andreas  
APPLICANT: Hirsch, David J.  
APPLICANT: Lodish, Harvey F.  
APPLICANT: Gimeno, Ruth E.  
APPLICANT: Tartaglia, Louis A.  
TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
FILE REFERENCE: WHI97-21P3MD  
CURRENT APPLICATION NUMBER: US/09/232,195A  
CURRENT FILING DATE: 1999-01-04

EARLIER APPLICATION NUMBER: 60/071,374  
EARLIER FILING DATE: 1998-01-15  
EARLIER APPLICATION NUMBER: 60/093,491  
EARLIER FILING DATE: 1998-07-20  
EARLIER APPLICATION NUMBER: 60/110,941  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 50  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-232-195-50

Query Match 0.8%; Score 22; DB 4; Length 1173;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 1140 TGACAAAAAAAAAAAAAAAAAAAA 1161

## RESULT 35

US-09-064-411A-35  
Sequence 35, Application US/09064411A  
Patent No. 6331664  
GENERAL INFORMATION:  
APPLICANT: Rubin-Wilson, Beth  
APPLICANT: Guo, Lining  
APPLICANT: Skokut, Tom  
APPLICANT: Young, Scott  
APPLICANT: Folkerts, Otto  
APPLICANT: Armstrong, Katherine  
APPLICANT: Cowen, Neil M.  
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP  
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and  
TITLE OF INVENTION: Their Use in The Regulation of Fatty Acid Content of Oil  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dow Agrosciences Patent Department  
STREET: 9330 Zionsville Road  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: USA  
ZIP: 46268  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/064,411A  
FILING DATE: 22-APR-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/045,827  
FILING DATE: 05-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Borucki, Andrea T.  
REGISTRATION NUMBER: 33651  
REFERENCE/DOCKET NUMBER: 50433  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-337-4846  
TELEFAX: 317-337-4847  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1287 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:



NAME/KEY: CDS  
LOCATION: 38.1135  
FEATURE:  
NAME/KEY: mat\_deptide  
LOCATION: 140.1135  
US-09-064-411A-35

Query Match 0.8%; Score 22; DB 3; Length 1287;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGCACAAAAAAAAAAAAAAAAA 2679  
Db 1265 TGCACAAAAAAAAAAAAAAAAA 1266

RESULT 36  
US-08-979-608A-10  
Sequence 10, Application US/08979608A  
Patent No. 6355451  
GENERAL INFORMATION:

APPLICANT: Lees, Ann M.  
Lees, Robert S.  
Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,608A  
FILING DATE: 26-Nov. 6355451-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/048,547  
FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930  
FILING DATE: 27-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Louis  
REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)  
TELEPHONE: 617/542-8906  
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:

NAME/KEY: Coding Sequence  
LOCATION: 58...510

SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-08-979-608A-10

Query Match 0.8%; Score 22; DB 3; Length 1404;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2678

Db 1383 TTGACAAAAAAAAAAAAAAAAA 1404

RESULT 37  
US-09-517-849-10  
Sequence 10, Application US/09517849  
Patent No. 6605588  
GENERAL INFORMATION:

APPLICANT: Lees, Ann M.  
Lees, Robert S.  
Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN  
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND  
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/517,849  
FILING DATE: 02-Mar-2000

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/979,608  
FILING DATE: 26-NOV-1997

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Louis  
REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: 10797-003001  
TELEPHONE: 617/542-8906  
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1404 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:

NAME/KEY: Coding Sequence  
LOCATION: 58...510

SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-517-849-10

Query Match 0.8%; Score 22; DB 4; Length 1404;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2678  
Db 1383 TTGACAAAAAAAAAAAAAAAAA 1404

RESULT 38

US-09-616-289-10  
Sequence 10, Application US/09616289  
Patent No. 6632923

GENERAL INFORMATION:

APPLICANT: Lees, Ann M.

APPLICANT: Lees, Robert S.

APPLICANT: Law, Simon W.

APPLICANT: Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING

```

; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (510)
US-09-616-289-10

Query Match          0.8%; Score 22; DB 4; Length 1404;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2678
DB 1383 TTGACAAAAAAAAAAAAAAAAA 1404

RESULT 39
US-09-347-650-15
; Sequence 15, Application US/09347650
; Patent No. 6576814
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven
; TITLE OF INVENTION: Manipulation of MLO Genes to Enhance Disease Resistance
; FILE REFERENCE: 5718-42035718/158714
; CURRENT APPLICATION NUMBER: US/09/347,650
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: MLO9
US-09-347-650-15

Query Match          0.8%; Score 22; DB 4; Length 1557;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAA 2679
DB 1531 TGACAAAAAAAAAAAAAAAAA 1552

RESULT 40
US-09-268-140-11
; Sequence 11, Application US/09268140
; Patent No. 6268176
; GENERAL INFORMATION:
; APPLICANT: Gemmill, Robert M.
; APPLICANT: Drabkin, Harry A.
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED
; FILE REFERENCE: 93445-00004
; CURRENT APPLICATION NUMBER: US/09/268,140
```

```

; CURRENT FILING DATE: 2000-03-12
; PRIOR APPLICATION NUMBER: US 60/077,723
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-268-140-11

Query Match          0.8%; Score 22; DB 3; Length 2339;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAA 2679
DB 2317 TGACAAAAAAAAAAAAAAAAA 2338

RESULT 41
US-07-867-106-5/c
; Sequence 5, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2422 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
US-07-867-106-5

Query Match          0.8%; Score 22; DB 1; Length 2422;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAA 2679
```

DB 2295 TGACAAAAAAAAAAAAAAAAAAAA 2274

## RESULT 42

US-09-268-140-1  
Sequence 1, Application US/09268140  
Patent No. 6268176  
GENERAL INFORMATION:  
APPLICANT: Gemmell, Robert M.  
TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED  
FILE REFERENCE: 93445-00004  
CURRENT APPLICATION NUMBER: US/09/268,140  
CURRENT FILING DATE: 2000-03-12  
PRIOR APPLICATION NUMBER: US 60/077,723  
PRIOR FILING DATE: 1998-03-12  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 2505  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (238) ..(2232)  
US-09-268-140-1

## Query Match

Best Local Similarity 0.8%; Score 22; DB 3; Length 2505;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 2483 TGACAAAAAAAAAAAAAAAAAAAA 2504

## RESULT 43

US-09-268-140-7  
Sequence 7, Application US/09268140  
Patent No. 6268176  
GENERAL INFORMATION:  
APPLICANT: Gemmell, Robert M.  
TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED  
FILE REFERENCE: 93445-00004  
CURRENT APPLICATION NUMBER: US/09/268,140  
CURRENT FILING DATE: 2000-03-12  
PRIOR APPLICATION NUMBER: US 60/077,723  
PRIOR FILING DATE: 1998-03-12  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 7  
LENGTH: 2517  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: repeat unit  
LOCATION: (165) ..(176)  
US-09-268-140-7

## Query Match

Best Local Similarity 0.8%; Score 22; DB 3; Length 2517;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 2495 TGACAAAAAAAAAAAAAAAAAAAA 2516

## RESULT 44

US-09-604-605-1  
Sequence 1, Application US/09604605  
Patent No. 6421613

## GENERAL INFORMATION:

APPLICANT: Ramgopal Nadimpalli  
APPLICANT: Carl R. Simons  
TITLE OF INVENTION: Maize Prolifera Gene and Uses Thereof  
FILE REFERENCE: 1138  
CURRENT APPLICATION NUMBER: US/09/604,605  
CURRENT FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/143,222  
PRIOR FILING DATE: 1999-07-09  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2533  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (110) ..(2269)  
US-09-604-605-1

## Query Match

Best Local Similarity 0.8%; Score 22; DB 4; Length 2533;  
Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 2509 TGACAAAAAAAAAAAAAAAAAAAA 2530

## RESULT 45

US-08-160-861-1  
Sequence 1, Application US/08160861  
Patent No. 5539295  
GENERAL INFORMATION:  
APPLICANT: MORRIS, STEPHAN W  
TITLE OF INVENTION: NOVEL FUSION NUCLEIC ACID SEQUENCES AND  
TITLE OF INVENTION: FUSION PROTEINS PRESENT IN HUMAN T(2;5) LYMPHOMA, METHODS  
TITLE OF INVENTION: OF DETECTION AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSER: STERNF, KESSLER, GOLDSTEIN & FOX  
STREET: 1100 NEW YORK AVE NW SUITE 600  
CITY: WASHINGTON  
STATE: D.C.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/160,861  
FILING DATE: 02-DEC-1993  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A  
REGISTRATION NUMBER: 36217  
REFERENCE/DOCKET NUMBER: 0656.0400000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2678  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2608 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1719  
US-08-160-861-1

## Query Match

Best Local Similarity 0.8%; Score 22; DB 1; Length 2608;

Best Local Similarity 100.0%; Pred. No. 9.7;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 2585 TGACAAAAAAAAAAAAAAAAAAAA 2606

## RESULT 46

US-10-140-002-345  
Sequence 345, Application US/10140002

Patent No. 6725730

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerltisen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P330R1C59

CURRENT APPLICATION NUMBER: US/10/140,002

CURRENT FILING DATE: 2002-05-06

Prior Application removed - See Palm or File Wrapper

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 345

LENGTH: 2933

TYPE: DNA

ORGANISM: Homo Sapien

US-10-140-002-345

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 2872 TGACAAAAAAAAAAAAAAAAAAAA 2893

## RESULT 47

US-07-867-106-4/C  
Sequence 4, Application US/07867106

Patent No. 5389526

GENERAL INFORMATION:

APPLICANT: Slade, Martin B

APPLICANT: Chang, Andy C M

APPLICANT: Williams, Keith L

TITLE OF INVENTION: Improved Plasmid Vectors for Cellular

TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSER: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris

STREET: One Liberty Place 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/867,106

FILING DATE: 19920625

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PJ 7187

APPLICATION NUMBER: PCT/AU90/00530

FILING DATE: 02-NOV-1989

ATTORNEY/AGENT INFORMATION:

NAME: Feeney, Joanne Longo

REGISTRATION NUMBER: 35,134

REFERENCE/DOCKET NUMBER: RICE-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3138 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

ANTI-SENSE: NO

US-07-867-106-4

QY 2658 TGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 3011 TGACAAAAAAAAAAAAAAAAAAAA 2990

## RESULT 48

US-07-867-106-2  
Sequence 2, Application US/07867106

Patent No. 5389526

GENERAL INFORMATION:

APPLICANT: Slade, Martin B

APPLICANT: Chang, Andy C M

APPLICANT: Williams, Keith L

TITLE OF INVENTION: Improved Plasmid Vectors for Cellular

TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSER: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris

STREET: One Liberty Place 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/867,106

FILING DATE: 19920625

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PJ 7187

APPLICATION NUMBER: PCT/AU90/00530

FILING DATE: 02-NOV-1989

ATTORNEY/AGENT INFORMATION:

NAME: Feeney, Joanne Longo

REGISTRATION NUMBER: 35,134

REFERENCE/DOCKET NUMBER: RICE-0002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 5852 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2378..5038  
NAME/KEY: CDS  
LOCATION: 2378..5038  
US-07-867-106-2

Query Match 0.8%; Score 22; DB 1; Length 5852;  
Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGCACAAAAAAAAAAAAAAAAA 2679  
DB 943 TGCACAAAAAAAAAAAAAAAAA 964

RESULT 49  
US-07-867-106-2/c  
Sequence 2, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526r1s  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867,106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
APPLICATION NUMBER: PCT/NU90/00530  
FILING DATE: 02-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Peeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5852 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2378..5038  
FEATURE:

NAME/KEY: CDS  
LOCATION: 2378..5038  
US-07-867-106-2

Query Match 0.8%; Score 22; DB 1; Length 5852;  
Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2658 TGCACAAAAAAAAAAAAAAAAA 2679  
DB 1647 TGCACAAAAAAAAAAAAAAAAA 1626

RESULT 50  
US-08-836-022A-10  
Sequence 10, Application US/08836022A  
Patent No. 6001557  
GENERAL INFORMATION:  
APPLICANT: Trustees of the University of Pennsylvania  
APPLICANT: Wilson, James M.  
APPLICANT: Fisher, Krishna J.  
APPLICANT: Chen, Shu-Jen  
APPLICANT: Weitzman, Matthew  
TITLE OF INVENTION: Improved Adenovirus Virus and  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/836,022A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/331,381  
FILING DATE: 28-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: GNPVN.008PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19307 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-08-836-022A-10

Query Match 0.8%; Score 22; DB 3; Length 19307;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAA 2678  
DB 14579 TTGACAAAAAAAAAAAAAAAAA 14600

Search completed: January 15, 2005, 12:28:41  
Job time : 253 secs

***This Page Blank (uspio)***



















962 22 0.8 1486 13 US-10-010-050A-1 Sequence 1, Appli  
 963 22 0.8 1489 18 US-10-739-930-2786 Sequence 2786, Ap  
 964 22 0.8 1574 17 US-10-437-963-64454 Sequence 64454, A  
 965 22 0.8 1592 18 US-10-416-330-5 Sequence 5, Appli  
 966 22 0.8 1597 16 US-10-287-971-99 Sequence 99, Appli  
 967 22 0.8 1606 16 US-10-264-237-664 Sequence 664, App  
 968 22 0.8 1635 18 US-10-425-115-142290 Sequence 142290,  
 969 22 0.8 1641 14 US-10-097-065-108 Sequence 108, App  
 970 22 0.8 1641 15 US-10-372-876-108 Sequence 108, App  
 971 22 0.8 1650 17 US-10-767-701-14815 Sequence 14815, A  
 972 22 0.8 1665 18 US-10-425-115-82818 Sequence 82818, A  
 973 22 0.8 1676 18 US-10-425-115-35855 Sequence 35855, A  
 974 22 0.8 1752 18 US-10-425-115-91391 Sequence 91391, A  
 975 22 0.8 1757 18 US-10-425-115-125815 Sequence 125815,  
 976 22 0.8 1768 10 US-09-997-003-20 Sequence 20, Appli  
 977 22 0.8 1795 18 US-10-425-115-127167 Sequence 127167,  
 978 22 0.8 1803 18 US-10-723-860-5943 Sequence 5943, Ap  
 979 22 0.8 1824 18 US-10-425-115-46816 Sequence 46816, A  
 980 22 0.8 1846 10 US-09-822-846-263 Sequence 263, App  
 981 22 0.8 1864 14 US-10-223-076-14 Sequence 14, Appli  
 982 22 0.8 1979 18 US-10-723-860-6704 Sequence 6704, Ap  
 983 22 0.8 1982 18 US-10-739-930-1030 Sequence 1030, Ap  
 984 22 0.8 1986 18 US-10-425-115-174166 Sequence 174166,  
 985 22 0.8 1987 16 US-10-264-237-115 Sequence 115, App  
 986 22 0.8 2000 9 US-09-938-842A-3674 Sequence 3674, Ap  
 987 22 0.8 2000 9 US-09-938-842A-4010 Sequence 4010, Ap  
 988 22 0.8 2000 9 US-09-938-842A-4017 Sequence 4017, Ap  
 989 22 0.8 2000 9 US-09-938-842A-4773 Sequence 4773, Ap  
 990 22 0.8 2000 9 US-09-938-842A-5013 Sequence 5013, Ap  
 991 22 0.8 2000 11 US-09-938-842A-3674 Sequence 3674, Ap  
 992 22 0.8 2000 11 US-09-938-842A-4010 Sequence 4010, Ap  
 993 22 0.8 2000 11 US-09-938-842A-4017 Sequence 4017, Ap  
 994 22 0.8 2000 11 US-09-938-842A-4773 Sequence 4773, Ap  
 995 22 0.8 2000 11 US-09-938-842A-5013 Sequence 5013, Ap  
 996 22 0.8 2074 18 US-10-425-115-120235 Sequence 120235,  
 997 22 0.8 2104 18 US-10-739-930-845 Sequence 845, App  
 998 22 0.8 2173 18 US-10-723-860-1685 Sequence 1685, Ap  
 999 22 0.8 2187 9 US-09-984-245-93 Sequence 93, Appli  
 1000 22 0.8 2187 10 US-09-966-262-93 Sequence 93, Appli

## ALIGNMENTS

RESULT 1  
 US-09-780-525-1 Application US/09780525  
 ; Sequence 1, Patent No. US20020004223A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bin-Bing Zhou  
 ; APPLICANT: Yuan Zhu  
 ; APPLICANT: Priya Chaturvedi  
 ; APPLICANT: Mark R. Hurle  
 ; APPLICANT: Xiaolong Li  
 ; TITLE OF INVENTION: PHARL, A NEW RING FINGER PROTEIN  
 ; FILE REFERENCE: GP-70668-C1  
 ; CURRENT APPLICATION NUMBER: US/09/780,525  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 09/456,876  
 ; PRIOR FILING DATE: 1999-12-08  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1995  
 ; TYPE: DNA  
 ; ORGANISM: HOMO SAPIENS  
 US-09-780-525-1

Query Match 74.5%; Score 1995; DB 9; Length 1995;  
 Best Local Similarity 100.0%; Pred. 0;  
 Matches 1995; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

91 ATGGAGCGCCGAGAGCAAGCAAGTGCAGCCGCGCCGAGCCCTGGGAGACGGCTCCG 150

Db 1 ATGGAGCGCCGAGAGCAAGCAAGTGCAGCCGCGCCGAGCCCTGGGAGACGGCTCCG 60  
 151 CGTCTGGCGCGAGAGAGAGCGAGCCGACGCTCTCTAGAGAAAGCGAGTGAACATC 210  
 61 CGTCTGGCGCGAGAGAGAGCGAGCCGACGCTCTCTAGAGAAAGCGAGTGAACATC 120  
 211 GGGCGGAGACGAGGTGGACCTTTCCTTCCCGAGCATTAATCTGTGTCTTGGAGATCAG 270  
 121 GGGCGGAGACGAGGTGGACCTTTCCTTCCCGAGCATTAATCTGTGTCTTGGAGATCAG 180  
 271 TGTGAATTTGTAGTGAATGAATAATCAGTCAAGGTGACATCTGAATACCGACCAAGT 330  
 181 TGTGAATTTGTAGTGAATGAATAATCAGTCAAGGTGACATCTGAATACCGACCAAGT 240  
 331 GGAACAGTGAATTAACAGCTGAAGGTGTGTAAGAGCAGACATGCCCTTTACAGCTGGG 390  
 241 GGAACAGTGAATTAACAGCTGAAGGTGTGTAAGAGCAGACATGCCCTTTACAGCTGGG 300  
 391 GATGTCATCTACTTGTGTGTAAGAGATGAATCCGGAACACACAGTGGCATTCCTAT 450  
 301 GATGTCATCTACTTGTGTGTAAGAGATGAATCCGGAACACACAGTGGCATTCCTAT 360  
 451 GAATCTTAATGAATAAGCAAGGATGACACAGAAATCCTTTGAAGCTTAACAGGAATAAT 510  
 361 GAATCTTAATGAATAAGCAAGGATGACACAGAAATCCTTTGAAGCTTAACAGGAATAAT 420  
 511 GTGTTCATGGAGCCAAAGATACCTCAGTGCAGAGTGCAGAGGCGAGGCGGATCCCGG 570  
 421 GTGTTCATGGAGCCAAAGATACCTCAGTGCAGAGTGCAGAGGCGAGGCGGATCCCGG 480  
 571 GTCCCTCGGTGTGCGCCGCACTCAGTGTGCTTTAGAGAACCAACCATCAATCG 630  
 481 GTCCCTCGGTGTGCGCCGCACTCAGTGTGCTTTAGAGAACCAACCATCAATCG 540  
 631 AGCTGAGACCTCTTCCCAAGCCTCGGCTCTTCCAGAGACCTTCTCTGACAGAGGCA 690  
 541 AGCTGAGACCTCTTCCCAAGCCTCGGCTCTTCCAGAGACCTTCTCTGACAGAGGCA 600  
 691 GAGCGTTCCTCAGTGTGAGGTCTGGGGGTGGTGCATCTCCCTTAAGAGAGTGTCC 750  
 601 GAGCGTTCCTCAGTGTGAGGTCTGGGGGTGGTGCATCTCCCTTAAGAGAGTGTCC 660  
 751 TCTGTGCAAGTGAAGTCTCCAGCTTCTTCCCTTCCAGAGAGAGAGTGTCC 810  
 661 TCTGTGCAAGTGAAGTCTCCAGCTTCTTCCCTTCCAGAGAGAGAGTGTCC 720  
 811 TCTTTTGTGTGAGAACCCAGATCAGAGATTTGAGAGCCGTGAAGAGAAATG 870  
 721 TCTTTTGTGTGAGAACCCAGATCAGAGATTTGAGAGCCGTGAAGAGAAATG 780  
 871 AGAGAGATGGAGACCTTGACCTGAGCGGCGAGTGTGTGCGCAACCCGCTGAAT 930  
 781 AGAGAGATGGAGACCTTGACCTGAGCGGCGAGTGTGTGCGCAACCCGCTGAAT 840  
 931 GCCCAACCGTCCAGAGAGATGAGACCGGTGGAGAGCCAGAGAGAGAGAGAG 990  
 841 GCCCAACCGTCCAGAGAGATGAGACCGGTGGAGAGCCAGAGAGAGAGAGAGAG 900  
 991 AGCTGACATGATCATCTGCGACAGACCTGCTGACAGCTGGGTGAGTGGAGCCCTGC 1050  
 901 AGCTGACATGATCATCTGCGACAGACCTGCTGACAGCTGGGTGAGTGGAGCCCTGC 960  
 1051 ATGCACACGTTCTGCGGCGCTTGTCTACTCGGCGTGGATGAGAGCGCTGCTGTGTCT 1110  
 961 ATGCACACGTTCTGCGGCGCTTGTCTACTCGGCGTGGATGAGAGCGCTGCTGTGTCT 1020  
 1111 ACTTCCGCTGTCTCCGTGAGAGCGATCTGTAAACCAATCTCAACACTCGTGGA 1170  
 1021 ACTTCCGCTGTCTCCGTGAGAGCGATCTGTAAACCAATCTCAACACTCGTGGA 1080  
 1171 GCATACCTCATTCAGATTCAGACAGAGATGAGAGTGTGAGAGAGTGTGAT 1230

Db 1081 GCATACCTCATCCAGATCCAGACAAGATGCGAGTGAAGAGATGTGAAAGTATGAT 1140  
QY 1231 GCAGAGATAAATCACTCAAGAATGCTGACGCCAAGTCAAGGCTTTTCTGAT 1290  
Db 1141 GCCAGAAATAAATCACTCAAGATGCTGACGCCAAGTCAAGGCTTTTCTGAT 1200  
QY 1291 GAAGAGAGATTCAGAGAGATGCTGAGCTGTCAAGCTTTGACAGTGAAGTCTCAGAC 1350  
Db 1201 GAAGAGAGATTCAGAGAGATGCTGAGCTGTCAAGCTTTGACAGTGAAGTCTCAGAC 1260  
QY 1351 ATTTAGCAGCATATGCTGTGTGCTGAGCTGTCTGAGTACAGAAAGCAGGCGCGCAG 1410  
Db 1261 ATTTAGCAGCATATGCTGTGTGCTGAGCTGTCTGAGTACAGAAAGCAGGCGCGCAG 1320  
QY 1411 CCTCCCATGAGCCAGACCCGAGGCGAGGCGAGGCGGAGGCCCAAGGCCCTTGGGGATGCA 1470  
Db 1321 CCTCCCATGAGCCAGACCCGAGGCGAGGCGAGGCGGAGGCCCAAGGCCCTTGGGGATGCA 1380  
QY 1471 CCTCCCATGAGCCAGACCCGAGGCGAGGCGAGGCGGAGGCCCAAGGCCCTTGGGGATGCA 1530  
Db 1381 CCTCCCATGAGCCAGACCCGAGGCGAGGCGAGGCGGAGGCCCAAGGCCCTTGGGGATGCA 1440  
QY 1531 AGCCACGCTGTGACCTGCTGCTTCCAGCCATGCCGACCGAGAGCGAGCGCGAG 1590  
Db 1441 AGCCACGCTGTGACCTGCTGCTTCCAGCCATGCCGACCGAGAGCGAGCGCGAG 1500  
QY 1591 CAGGACCGGCTGTGACCTGCTGCTTCCAGCCATGCCGACCGAGAGCGAGCGCGAG 1650  
Db 1501 CAGGACCGGCTGTGACCTGCTGCTTCCAGCCATGCCGACCGAGAGCGAGCGCGAG 1560  
QY 1651 TACTGGGGCTGACCCGAGCCGCTGCTACGAGCTGCTGAGCCCGTTTGTGAGCTCAAC 1710  
Db 1561 TACTGGGGCTGACCCGAGCCGCTGCTACGAGCTGCTGAGCCCGTTTGTGAGCTCAAC 1620  
QY 1711 CTGGGTGACAGTGTGTGACCGGCTGTGACCAAGCAAGCTTCAAGCATCTG 1770  
Db 1621 CTGGGTGACAGTGTGTGACCGGCTGTGACCAAGCAAGCTTCAAGCATCTG 1680  
QY 1771 AAGAAATTAATCTGACCAAGAGTTTGAATGAAATTAATGTAACCGAGCCCTG 1830  
Db 1681 AAGAAATTAATCTGACCAAGAGTTTGAATGAAATTAATGTAACCGAGCCCTG 1740  
QY 1831 GCTCTCAGAGGAGATGTTTCTGCTGTCTGATTAAGAGTCAAGGAGCAACCTG 1890  
Db 1741 GCTCTCAGAGGAGATGTTTCTGCTGTCTGATTAAGAGTCAAGGAGCAACCTG 1800  
QY 1891 TGTTACTGTGTGCTGTGAGCTTCCGTGAGCTGACCTATCAGTATCGGAGACAT 1950  
Db 1801 TGTTACTGTGTGCTGTGAGCTTCCGTGAGCTGACCTATCAGTATCGGAGACAT 1860  
QY 1951 CCTGCTTCGAGTGTGACGCTGCTGATCAATCCGCTGCTGATCTGAGGCGCTTAC 2010  
Db 1861 CCTGCTTCGAGTGTGACGCTGCTGATCAATCCGCTGCTGATCTGAGGCGCTTAC 1920  
QY 2011 TSCCGACTCAGGTGAAGCTCAGCAGGCGATGAATTAATCATATCTGTGAAGAGA 2070  
Db 1921 TSCCGACTCAGGTGAAGCTCAGCAGGCGATGAATTAATCATATCTGTGAAGAGA 1980  
QY 2071 AGTTCAAAACTAA 2085  
Db 1981 AGTTCAAAACTAA 1995

## RESULT 2

US-10-108-260A-2073  
; Sequence 2073, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27

; NUMBER OF SEQ ID NOS: 548  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO: 2073  
; LENGTH: 2448  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-2073  
Query Match 35.2%; Score 944; DB 16; Length 2448;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 994; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 708 TGGGTCTGGGGTGTGGATCTCCCTAAAGAAAGTGTCTGTGCAAGTATGA 767  
Db 357 TGGGTCTGGGGTGTGGATCTCCCTAAAGAAAGTGTCTGTGCAAGTATGA 416  
QY 768 AGTCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGTGTCTTTGTGCTTGA 827  
Db 417 AGTCTCAGCTTTGCTCAGCTCTCCAGACAGAAAGTGTCTTTGTGCTTGA 476  
QY 828 ACCCCAGATCAGAGAGATTTGAGCCCTGAAAGAAATGAGAGAGATGGGAGCT 887  
Db 477 ACCCCAGATCAGAGAGATTTGAGCCCTGAAAGAAATGAGAGAGATGGGAGCT 536  
QY 888 TGACCTGAACGGGCAAGTTGTTGTGCAACACCGGTAAAGTCCCAACGCTCAGCA 947  
Db 537 TGACCTGAACGGGCAAGTTGTTGTGCAACACCGGTAAAGTCCCAACGCTCAGCA 596  
QY 948 GGAAGCTCAGAGAGCGGCTGGAGAGCAGACAAATGAGAGAGCTGACATGATCAT 1007  
Db 597 GGAAGCTCAGAGAGCGGCTGGAGAGCAGACAAATGAGAGAGCTGACATGATCAT 656  
QY 1008 CTGCCAGACCTGTGCTGACAGCTGCTGAGATTGCAAGCTCTGATGACATGCTGCGC 1067  
Db 657 CTGCCAGACCTGTGCTGACAGCTGCTGAGATTGCAAGCTCTGATGACATGCTGCGC 716  
QY 1068 GGCTTGCTATCTGGGGTGTGATGAGAGCTGCTCCGTGCTTACCTGCGCTGCTCCGT 1127  
Db 717 GGCTTGCTATCTGGGGTGTGATGAGAGCTGCTCCGTGCTTACCTGCGCTGCTCCGT 776  
QY 1128 GGAAGGAGATCTGTAAACCAATCTCTCAACATCTGTGAGAGATTAATCTATCCAGCA 1187  
Db 777 GGAAGGAGATCTGTAAACCAATCTCTCAACATCTGTGAGAGATTAATCTATCCAGCA 836  
QY 1188 TCCAGACAAAGATCGAGTGAAGAAAGTGTGCAAGTATGATGCCAGAAATTAATC 1247  
Db 837 TCCAGACAAAGATCGAGTGAAGAAAGTGTGCAAGTATGATGCCAGAAATTAATC 896  
QY 1248 TCAAGACATGTGCAAGCCCAAGTCAAGCGGCTTTTCTGATGAAGAGAGTTCA 1307  
Db 897 TCAAGACATGTGCAAGCCCAAGTCAAGCGGCTTTTCTGATGAAGAGAGTTCA 956  
QY 1308 GGAAGCTGTGAGCTGTGACAGTGAAGTGAAGTGTGATGATTAAGTGAAGTGAAGT 1367  
Db 957 GGAAGCTGTGAGCTGTGACAGTGAAGTGAAGTGTGATGATTAAGTGAAGTGAAGT 1016  
QY 1368 CGTGTGCGGAGATGTCTTGAATCAAGAGGCGGCGAGCTCTCCCATGCGCCAGC 1427  
Db 1017 CGTGTGCGGAGATGTCTTGAATCAAGAGGCGGCGAGCTCTCCCATGCGCCAGC 1076  
QY 1428 ACCCGAGGCGGAGCGAGAGCCCAAGGCGCTGGGGAGATGACCTTCAAGCTTGTGAG 1487  
Db 1077 ACCCGAGGCGGAGCGAGAGCCCAAGGCGCTGGGGAGATGACCTTCAAGCTTGTGAG 1136  
QY 1488 CTTGACGACAGAGTCCAGATTAAGTGTGCTCTGCAAGAGAACCAAGCTGTGAC 1547  
Db 1137 CTTGACGACAGAGTCCAGATTAAGTGTGCTCTGCAAGAGAACCAAGCTGTGAC 1196  
QY 1548 CTGCTGCTTCCAGCCCATGCGGAGCGAGAGCGGAGCGAGCAACCGCGTGTGCG 1607  
Db 1197 CTGCTGCTTCCAGCCCATGCGGAGCGGAGAGCGGAGCGGAGCAACCGCGTGTGCG 1256  
QY 1608 CCTCAGAGAGTGTGGGTCTGCTGCAAGCTTTTCTGCCACCTGTACTGGGCTGACCGG 1667

```

Db      1257 CCTGAGCAGTGTGGGTCCTGCTGAGCCCTTCTGCACTGTGGGCTGACCCG 1316
Qy      1668 GACCGGCTCTACGCGTGTGCTGCGCCCGCTTTGTG 1702
Db      1317 GACCGGCTCTACGCGTGTGCTGCGCCCGCTTTGTG 1351

RESULT 3
US-09-764-864-22
; Sequence 22, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1311
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1281)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-22

Query Match      28.6%; Score 765; DB 9; Length 1311;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 765; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      710 GGTCTGGGGGTGGTGGCATCTCCCTAAAGAAAGAGTGTGCTGCGAAGTGAAG 769
Db      383 GGTCTGGGGGTGGTGGCATCTCCCTAAAGAAAGTGTGCTGCGAAGTGAAG 442
Qy      770 TCTCCAGCTTGTGCTGCTGCTCTCCAGAGAAAGATCGTCTTTTGTGCTGGAA 829
Db      443 TCTCCAGCTTGTGCTGCTGCTCTCCAGAGAAAGATCGTCTTTTGTGCTGGAA 502
Qy      830 CCCAGATCAGAGAGATTTGAGAGCCCGTGAAGAAATGAGAGAGATGGGACCTTG 889
Db      503 CCCAGATCAGAGAGATTTGAGAGCCCGTGAAGAAATGAGAGAGATGGGACCTTG 562
Qy      890 ACCGTGAAGGGGAGTGTGTGTGCAACAACGGGTGAATGCGCAACCGTCCAGAG 949
Db      563 ACCGTGAAGGGGAGTGTGTGTGCAACAACGGGTGAATGCGCAACCGTCCAGAG 622
Qy      950 AGCTCAGACAGCGGCTGGGAACCAAGACAGATGAGAGAGCGTGAATGATCATCT 1009
Db      623 AGCTCAGACAGCGGCTGGGAACCAAGACAGATGAGAGAGCGTGAATGATCATCT 682
Qy      1010 GCCAGAGCTGTGCTGCACTGCTGTGATTTGAGAGCCCTGATGCAAGTTTGTGCGG 1069
Db      683 GCCAGAGCTGTGCTGCACTGCTGTGATTTGAGAGCCCTGATGCAAGTTTGTGCGG 742
Qy      1070 CTGTGCTACTCGGGCTGAGATGAGAGCGTGTGCTGTGCTGCAACCTGCGTGCAGG 1129
Db      743 CTGTGCTACTCGGGCTGAGATGAGAGCGTGTGCTGTGCTGCAACCTGCGTGCAGG 802
Qy      1130 AGCGATCTGTAAAAACAATCTCAACAACCTGTGAGAGATCACTCATCAACATC 1189
Db      803 AGCGATCTGTAAAAACAATCTCAACAACCTGTGAGAGATCACTCATCAACATC 862
Qy      1190 CAGACAAAGTCCGAGTGAAGAGATGTGAAGATTTGATGCGCAGAGATTAATCACTC 1249
Db      863 CAGACAAAGTCCGAGTGAAGAGATGTGAAGATTTGATGCGCAGAGATTAATCACTC 922
Qy      1250 AAGACATGCTGCAAGCCCAAGTCAAGGCGTCTTTCTGATGAAGAGGAGTTCAAGG 1309

```

```

Db      923 AAGACATGCTGCAAGCCCAAGTCAAGGCGTCTTTTCTGATGAAGAGAGATTCAAGG 982
Qy      1310 ACTGTCTGAGCTGTGACAGCTTGAACATGATCTTCAAGATTTAGCCATTCGTCG 1369
Db      983 ACTGTCTGAGCTGTGACAGCTTGAACATGATGATCTTCAAGATTTAGCCATTCGTCG 1042
Qy      1370 TGTGCGGCGAGTGTCTGTGATGAAGAGGCGGCGGCGAGCTGCCCATGCGCAGAGC 1429
Db      1043 TGTGCGGCGAGTGTCTGTGATGAAGAGGCGGCGGCGAGCTGCCCATGCGCAGAGC 1102
Qy      1430 CCGAGGCGGAGCCAGAGAGCCCAAGGCGCTGTGGGAGATGACCTT 1474
Db      1103 CCGAGGCGGAGCCAGAGAGCCCAAGGCGCTGTGGGAGATGACCTT 1147

RESULT 4
US-10-094-749-160
; Sequence 160, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUYU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: OTSUKA, KAORU
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 160
; LENGTH: 2186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-160

Query Match      23.6%; Score 633; DB 15; Length 2186;
Best Local Similarity 99.7%; Pred. No. 2,76-310;
Matches 863; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Qy      1796 TGACATGGAAGAAACATGTTGACCGAGAGCTGTGAGCTCTCCAGCGGGAGTGTTCCTC 1855
Db      695 TGACATGGAAGAAACATGTTGACCGAGAGCTGTGAGCTCTCCAGCGGGAGTGTTCCTC 754
Qy      1856 TGTCTGATTTACAGATGCAAGGAGACACCTTCTGTGTTTACGTCGTGCGCCAGCT 1915
Db      755 TGTCTGATTTACAGATGCAAGGAGACACCTTCTGTGTTTACGTCGTGCGCCAGCT 814
Qy      1916 TCCGTGAGTGAAGTATGATGATGCGAGAAATCTGCTTCCAGATTGCGAGTGGCCG 1975
Db      815 TCCGTGAGTGAAGTATGATGATGCGAGAAATCTGCTTCCAGATTGCGAGTGGCCG 874
Qy      1976 TAAACATCCCGTCTGATCTGATCTGAGGCGGTGTAAGTCCGCACTAGTGAAGCTCACC 2035
Db      875 TAAACATCCCGTCTGATCTGATCTGAGGCGGTGTAAGTCCGCACTAGTGAAGCTCACC 934

```



[illegible]

Qy	245	GCAATPAACTGGTCTCTGAGATCACTGTAGATTGTAGTGGATGATAAAATCAGGTCAGG	304
Db	51	GCAATPAACTGGTCTCTGAGATCACTGTAGATTGTAGTGGATGATAAAATCAGGTCAGG	110
Qy	305	TGACACCTGGAAAGATCCAGACCCCAATGGAACAGATTAACAAGCTTGAAGCTTTTAAGA	364
Db	111	TGACACCTGGAAAGATCCAGACCCCAATGGAACAGATTAACAAGCTTGAAGCTTTTAAGA	170
Qy	365	AGCAGACATGCGCTTTACAGATCGAGCTGGGAGTGTCACTACTTGGTGTACAGAAATGAAC	424
Db	171	AGCAGACATGCGCTTTACAGATCGAGCTGGGAGTGTCACTACTTGGTGTACAGAAATGAAC	230
Qy	425	CGGAACACAGAGTGGCATACCTCTATGTAATCTTTAAGTGAAGCAAGGCATGACACA	482

Query Match 8.9%; Score 238; DB 9; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 1e-109;  
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0

US-09-764-864-323  
 Sequence 323: Application US/09764864  
 Patent No. US20020132753A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PT23  
 CURRENT APPLICATION NUMBER: US/09/764,864  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 1792  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 323  
 LENGTH: 354  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURES:  
 NAME/KEY: SITE  
 LOCATION: (289)  
 OTHER INFORMATION: n equals a,t,g, or c  
 NAME/KEY: SITE  
 LOCATION: (313)  
 OTHER INFORMATION: n equals a,t,g, or c  
 US-09-764-864-323

```

Db          231 CGAACAACAAGTGGCACTACTATGATCTTTAAGTGAAGAAGCAAGCATGACACA 288

RESULT 8
US-09-764-864-742
: Sequence 742, Application US/09764864
: Patent No. US20020132753A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT23
: CURRENT APPLICATION NUMBER: US/09/764,864
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 742
: LENGTH: 354
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURES:
: NAME/KEY: SITE
: LOCATION: (289)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (313)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-742

```

Query Match	8.9%;	Score 238;	DB 9;	Length 354;
Best Local Similarity	100.0%;	Prod. No. 1e-109;		
Matches 238;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

  

Qy	245	GCATTAACCTGCTCTCTGGAGATCACTGTAGATTGTAGTGAGTGAATAATCAGGTCAGG	304
Db	51	GCATTAACCTGCTCTCTGGAGATCACTGTAGATTGTAGTGAGTGAATAATCAGGTCAGG	110
Qy	305	TGACACTGGAAGATPACAGCACACAGTGGAAAGTGAATTAACAAGCTGAAGTGTTAAGA	364
Db	111	TGACACTGGAAGATPACAGCACACAGTGGAAAGTGAATTAACAAGCTGAAGTGTTAAGA	170
Qy	365	AGCAGACATGCCCTTTACAGACTGGGGATGTCATCTACTTGGTGAACGAGAATGAAC	424
Db	171	AGCAGACATGCCCTTTACAGACTGGGGATGTCATCTACTTGGTGAACGAGAATGAAC	230
Qy	425	CGAACAACAAGTGGCATCACTCTATGAACTTTTAAAGTGAAGAAGCAAGCATGACACA	482
Db	231	CGAACAACAAGTGGCATCACTCTATGAACTTTTAAAGTGAAGAAGCAAGCATGACACA	288

  

RESULT 9
US-09-918-995-13662
Sequence 13662, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 13662
LENGTH: 449
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc-feature
LOCATION: (1)...(449)
OTHER INFORMATION: n = A,T,C or G



US-09-918-995-13662

Query Match  
Best Local Similarity 100.0%; Pred. No. 4.6e-99;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2445 AGTTTAAAGGAGAGAGGCAAGTTTCAAAAACATTTGTTTCAGAGAGGAGCATTAAG 2504  
DB 92 AGTTTCAAGGAGAGAGGCAAGTTTCAAAAACATTTGTTTCAGAGAGGAGCATTAAG 151  
QY 2505 TTTTACAGCTTACAGAGAGTACACAAATATCTGCTGGGAAAACACAGCATTTATCT 2564  
DB 152 TTTTACAGCTTACAGAGAGTACACAAATATCTGCTGGGAAAACACAGCATTTATCT 211  
QY 2565 ATTTTATTTTATAGTTTGGTGTATCTCTATATAGATTTTAAATGTCACAAATG 2624  
DB 212 ATTTTATTTTATAGTTTGGTGTATCTCTATATAGATTTTAAATGTCACAAATG 271  
QY 2625 TAGCACAATAATATATATTTATATTTACAAATTGAC 2661  
DB 272 TAGCACAATAATATATATTTATATTTACAAATTGAC 308

RESULT 10

US-09-918-995-2180  
; Sequence 2180, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: HySeq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2180  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(476)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-2180

Query Match  
Best Local Similarity 100.0%; Pred. No. 5.6e-28;  
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2046 ATTCAATCATATCTGTGAACAGACAAAGTTCAAAAACTAAGATCCAGAGCCCTGAGCA 2105  
DB 52 ATTCAATCATATCTGTGAACAGACAAAGTTCAAAAACTAAGATCCAGAGCCCTGAGCA 111  
QY 2106 GCTTTACGACCTGAGG 2122  
DB 112 GCTTTACGACCTGAGG 128

RESULT 11

US-09-908-975-6220  
; Sequence 6220, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: FAIGER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING  
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME

; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6220  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-908-975-6220

Query Match  
Best Local Similarity 100.0%; Pred. No. 2.5e-19;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1843 GGAGTGTCTGCTGTCTGATTACAGAGTCACGGGAGACCGGTTCTGTACTGCTGT 1902  
DB 1 GGAGTGTCTGCTGTCTGATTACAGAGTCACGGGAGACCGGTTCTGTACTGCTGT 60

RESULT 12

US-10-425-115-112922/C  
; Sequence 112922, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 112922  
; LENGTH: 281  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_34473C.1  
US-10-425-115-112922

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.045;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2654 AAATTGACAAAAA 2679  
DB 43 AAATTGACAAAAA 18

RESULT 13

US-10-719-993-23863/C  
; Sequence 23863, Application US/10719993  
; Publication No. US20040265849A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001496  
; CURRENT APPLICATION NUMBER: US/10/719,993  
; CURRENT FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 55342  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23863  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-719-993-23863

Query Match 0.9%; Score 25; DB 18; Length 201;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAAAAAAAAAAAAAAAAA 2679

DB 26 AATTGACAAAAAAAAAAAAAAAAAAAA 2

RESULT 14

US-10-304-928-15  
; Sequence 15, Application US/10304928  
; Publication No. US20030101484A1  
; GENERAL INFORMATION:  
; APPLICANT: Yalpani, Carl R.  
; APPLICANT: Yalpani, Nasser  
; TITLE OF INVENTION: Maize Chitinases and Their Use in  
; FILE REFERENCE: 1100D  
; CURRENT APPLICATION NUMBER: US/10/304,928  
; PRIOR FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/522,714  
; PRIOR FILING DATE: 2000-03-10  
; PRIOR APPLICATION NUMBER: 60/125,915  
; PRIOR FILING DATE: 1999-03-24  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 15  
; LENGTH: 1085  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (46)...(966)  
US-10-304-928-15

Query Match 0.9%; Score 25; DB 15; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAAAAAAAAAAAAAAAAA 2679

DB 1061 AATTGACAAAAAAAAAAAAAAAAAAAA 1085

RESULT 15

US-10-425-115-124808  
; Sequence 124808, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 124808  
; LENGTH: 1253  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: M74577\_45300C.1  
US-10-425-115-124808

Query Match 0.9%; Score 25; DB 18; Length 1253;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAAAAAAAAAAAAAAAAA 2679  
DB 1197 AATTGACAAAAAAAAAAAAAAAAAAAA 1221

RESULT 16

US-09-957-664-1  
; Sequence 1, Application US/09957664  
; Patent No. US20020123097A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A. J.  
; TITLE OF INVENTION: 63760, A NOVEL HUMAN TRANSPORTER AND  
; FILE REFERENCE: US/09/957,664  
; CURRENT APPLICATION NUMBER: US/09/957,664  
; CURRENT FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/233,790  
; PRIOR FILING DATE: 2000-09-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1963  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (67)...(1491)  
US-09-957-664-1

Query Match 0.9%; Score 25; DB 9; Length 1963;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAAAAAAAAAAAAAAAAA 2679

DB 1916 AATTGACAAAAAAAAAAAAAAAAAAAA 1940

RESULT 17

US-10-154-419-15  
; Sequence 15, Application US/10154419  
; Publication No. US20030143675A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Gluckman, Maria Alexandra  
; APPLICANT: Meyers, Rachel E.  
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,  
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,  
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALFL, FBH58295FL, 57255,  
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-249  
; CURRENT APPLICATION NUMBER: US/10/154,419  
; CURRENT FILING DATE: 2002-05-22  
; PRIOR APPLICATION removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: FastSeq Version 4.0  
; SEQ ID NO 15  
; LENGTH: 1963  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (67)...(1491)  
US-10-154-419-15

Query Match 0.9%; Score 25; DB 15; Length 1963;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAAAAAAAAAAAAAAAAA 2679

DB 1916 AATTGACAAAAAAAAAAAAAAAAAAAA 1940

RESULT 18  
US-09-938-842A-3337/c  
; Sequence 3337, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 3337  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-3337

Query Match 0.9%; Score 25; DB 9; Length 2000;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679  
|||||  
1268 AATTGACAAAAA 1244  
|||||  
RESULT 19  
US-09-938-842A-3337/c  
; Sequence 3337, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 3337  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-3337

Query Match 0.9%; Score 25; DB 11; Length 2000;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679  
|||||  
1268 AATTGACAAAAA 1244  
|||||

RESULT 20  
US-10-719-993-6817/c  
; Sequence 6817, Application US/10719993  
; Publication No. US20040265849A1  
; GENERAL INFORMATION:  
; APPLICANT: Cargill, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001496  
; CURRENT APPLICATION NUMBER: US/10/719,993  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 55342  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6817  
; LENGTH: 717651  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(717651)  
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-  
US-10-719-993-6817

Query Match 0.9%; Score 25; DB 18; Length 717651;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2655 AATTGACAAAAA 2679  
|||||  
54671 AATTGACAAAAA 54647  
|||||  
RESULT 21  
US-09-867-701-9622  
; Sequence 9622, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; PRIOR FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9622  
; LENGTH: 199  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-9622

Query Match 0.9%; Score 24; DB 9; Length 199;  
Best Local Similarity 100.0%; Pred. No. 0.46;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 AATTGACAAAAA 2679  
|||||  
160 AATTGACAAAAA 183  
|||||

RESULT 22  
US-09-764-846-50  
; Sequence 50, Application US/09764846  
; Patent No. US20020102638A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT212  
; CURRENT APPLICATION NUMBER: US/09/764,846  
; PRIOR FILING DATE: 2001-01-17

```
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 348
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO: 50
;; LENGTH: 222
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (11)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (219)
;; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-846-50

Query Match          0.9%; Score 24; DB 9; Length 222;
Best Local Similarity 100.0%; Pred.No. 0.46;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      145 ATTGACAAAAAAAAAAAAAAAAAAAA 168

RESULT 23
US-09-764-846-122
;; Sequence 122, Application US/09764846
;; Patent No. US20020102638A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT212
;; CURRENT APPLICATION NUMBER: US/09/764,846
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 348
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO: 122
;; LENGTH: 222
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (11)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (219)
;; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-846-122

Query Match          0.9%; Score 24; DB 9; Length 222;
Best Local Similarity 100.0%; Pred.No. 0.46;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      145 ATTGACAAAAAAAAAAAAAAAAAAAA 168

RESULT 24
US-10-091-483-50
;; Sequence 50, Application US/10091483
;; Publication No. US20030049650A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT212C1
;; CURRENT APPLICATION NUMBER: US/10/091,483
;; CURRENT FILING DATE: 2002-03-07
;; NUMBER OF SEQ ID NOS: 348
;; Prior application removed - See File Wrapper or Palm
;; SOFTWARE: Patentln Ver. 2.0
```

```
;; SEQ ID NO: 50
;; LENGTH: 222
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (11)_feature
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (219)
;; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-483-50

Query Match          0.9%; Score 24; DB 14; Length 222;
Best Local Similarity 100.0%; Pred.No. 0.46;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      145 ATTGACAAAAAAAAAAAAAAAAAAAA 168

RESULT 25
US-10-091-483-122
;; Sequence 122, Application US/10091483
;; Publication No. US20030049650A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT212C1
;; CURRENT APPLICATION NUMBER: US/10/091,483
;; CURRENT FILING DATE: 2002-03-07
;; NUMBER OF SEQ ID NOS: 348
;; Prior application removed - See File Wrapper or Palm
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO: 122
;; LENGTH: 222
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (11)_feature
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (219)
;; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-483-122

Query Match          0.9%; Score 24; DB 14; Length 222;
Best Local Similarity 100.0%; Pred.No. 0.46;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      145 ATTGACAAAAAAAAAAAAAAAAAAAA 168

RESULT 26
US-10-357-930-20027
;; Sequence 20027, Application US/10357930
;; Publication No. US20040259086A1
;; GENERAL INFORMATION:
;; APPLICANT: Schlegel, Robert
;; APPLICANT: Endegge, Wilison
;; APPLICANT: Monahan, John
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;; TITLE OF INVENTION: HUMAN PROSTATS CANCER
;; FILE REFERENCE: MRI-007BCN
;; CURRENT APPLICATION NUMBER: US/10/357,930
;; CURRENT FILING DATE: 2003-02-04
;; PRIOR APPLICATION NUMBER: 09/785,276
;; PRIOR FILING DATE: 2003-02-16
```

```

; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20027
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2..4
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-20027
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 18; Length 291;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2656 ATTGACCAAAAAAAAAAAAAAAAAAAAA 2679
DB 199 ATTGACCAAAAAAAAAAAAAAAAAAAAA 222
```

```

RESULT 27
US-10-242-535A-21114
; Sequence 21114, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21114
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-21114
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 16; Length 320;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2656 ATTGACCAAAAAAAAAAAAAAAAAAAAA 2679
DB 289 ATTGACCAAAAAAAAAAAAAAAAAAAAA 312
```

```

RESULT 28
US-10-085-783A-21114
; Sequence 21114, Application US/10085783A
; Publication No. US20040037841A1
```

```

; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21114
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-21114
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 16; Length 320;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2656 ATTGACCAAAAAAAAAAAAAAAAAAAAA 2679
DB 289 ATTGACCAAAAAAAAAAAAAAAAAAAAA 312
```

```

RESULT 29
US-10-357-930-49814
; Sequence 49814, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlengel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49814
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 328..329..330..331..332..333..334..335..336..349..350..351
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-49814
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 18; Length 413;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

```

RESULT 30
US-09-918-995-10672
? Sequence 10672, Application US/09918995
? Publication NO. US20030073623A1
? GENERAL INFORMATION:
? APPLICANT: Hyseq, Inc.
? TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
? FROM VARIOUS CDNA LIBRARIES
? FILE REFERENCE: 20411-756
? CURRENT APPLICATION NUMBER: US/09/918,995
? PRIOR FILING DATE: 2001-07-30
? PRIOR APPLICATION NUMBER: US/09/235,076
? PRIOR FILING DATE: 1999-01-20
? NUMBER OF SEQ ID NOS: 38054
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 10672
? LENGTH: 474
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: mlec.feature
? LOCATION: (1)...(474)
? OTHER INFORMATION: n = A,T,C or G
? US-09-918-995-10672

```

		0.9%	Score	24	DB	10	Length	474	
Query Match Similarity		100.0%	Pred.	No.	0.46				
Best Local Similarity									
Matches	24;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
Qy	2656	ATTGACCAAAAAAAAAAAAAA		2679					
Dn	449	ATTGCAGCAAAAAAAAAAAAAAA		472					

```

RESULT 31
US-10-357-930-60931/c
: Sequence 60931, Application US/10357930
: Publication NO. US20040259086A1
: GENERAL INFORMATION:
: APPLICANT: Schlegel, Robert
: APPLICANT: Endege, Wilson
: APPLICANT: Monahan, John
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
: TITLE OF INVENTION: HUMAN PROSTATE CANCER
: FILE REFERENCE: MFI-00/BCN
: CURRENT APPLICATION NUMBER: US/10/357,930
: CURRENT FILING DATE: 2003-02-04
: PRIOR APPLICATION NUMBER: 09/785,276
: PRIOR FILING DATE: 2003-02-16
: PRIOR APPLICATION NUMBER: 60/183,319
: PRIOR FILING DATE: 2000-02-17
: PRIOR APPLICATION NUMBER: 60/189,862
: PRIOR FILING DATE: 2000-03-16
: PRIOR APPLICATION NUMBER: 60/207,454
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: 60/211,314
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/219,007
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/255,281
: PRIOR FILING DATE: 2000-12-13
: NUMBER OF SEQ ID NOS: 6232
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 60931
:
: TYPE: DNA
:

```

ORGANISM: Homo sapiens  
US-10-357-930-60931

Query Match	0.9%;	Score 24;	DB 18;	Length 525;
Best Local Similarity	100.0%;	Pred. No. 0.46;		
Matches 24;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 32
US-10-425-115-129128
; Sequence 129128, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 129128
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_49238C.1
US-10-425-115-129128

```

Query Match	0.9%;	Score 24;	DB 18;	length 566;
Best Local Similarity	100.0%;	Pred. No. 0.46;		
Matches 24;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

RESULT 33
US-10-027-632-147186/c
Sequence 147186, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185,218
PRIORITY FILING DATE: 2000-02-24
PRIORITY APPLICATION NUMBER: US 60/167,363
PRIORITY FILING DATE: 1999-11-23
PRIORITY APPLICATION NUMBER: US 60/156,358
PRIORITY FILING DATE: 1999-09-28
PRIORITY APPLICATION NUMBER: US 60/146,002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 147186
LENGTH: 635

```



TYPE: DNA  
ORGANISM: Human  
US-10-027-632-147186

Query Match  
Best Local Similarity 100.0%; Score 24; DB 13; Length 635;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2260 CCTCCCGAGAGCCCGGGGCGG 2283  
DB 315 CCTCCCGAGAGCCCGGGGCGG 292

## RESULT 34

US-10-027-632-147186/c  
Sequence 147186, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
PRIOR FILING DATE: 2000-07-12  
PRIOR FILING DATE: 2000-04-20  
PRIOR FILING DATE: 2000-03-29  
PRIOR FILING DATE: 2000-07-12  
PRIOR FILING DATE: 2000-02-24  
PRIOR FILING DATE: 1999-11-23  
PRIOR FILING DATE: 1999-09-28  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
LENGTH: 635  
SEQ ID NO 147186  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-147186

Query Match  
Best Local Similarity 100.0%; Score 24; DB 15; Length 635;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2260 CCTCCCGAGAGCCCGGGGCGG 2283  
DB 315 CCTCCCGAGAGCCCGGGGCGG 292

## RESULT 35

US-10-425-115-124214  
Sequence 124214, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 124214  
LENGTH: 677  
TYPE: DNA

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_44763C.1  
US-10-425-115-124214

Query Match  
Best Local Similarity 100.0%; Score 24; DB 18; Length 677;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACMAAAAAAAAAAAAAA 2679  
DB 647 ATTGACMAAAAAAAAAAAAAA 670

## RESULT 36

US-10-425-115-114433  
Sequence 114433, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 114433  
LENGTH: 688  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_35847C.1  
US-10-425-115-114433

Query Match  
Best Local Similarity 100.0%; Score 24; DB 18; Length 688;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACMAAAAAAAAAAAAAA 2679  
DB 652 ATTGACMAAAAAAAAAAAAAA 675

## RESULT 37

US-09-969-034-1012/c  
Sequence 1012, Application US/09969034  
Publication No. US20040110668A1  
GENERAL INFORMATION:  
APPLICANT: Burgess, Christopher C.  
APPLICANT: Aetle, Jon H.  
APPLICANT: Carroll, Eddie III  
APPLICANT: Catino, Theodore J.  
APPLICANT: Divedi, Poorima  
APPLICANT: Molino, Gary A.  
APPLICANT: Thiagalingam, Arunthathi  
APPLICANT: Lewis, Marcia B.  
TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
FILE REFERENCE: 1657/1032  
CURRENT FILING DATE: 2001-10-02  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 4494  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1012  
LENGTH: 703  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: misc feature  
LOCATION: 185, 257, 315, 326, 374, 379, 403, 405, 432, 442, 447, 480,  
LOCATION: 480, 492, 494, 497, 515, 518, 523, 556, 562, 563, 568, 593,  
LOCATION: 601, 607, 612, 613, 626, 627, 631, 634, 638, 641, 645, 654,  
LOCATION: 662, 672, 677, 696, 697, 703  
OTHER INFORMATION: n = A,T,C or G  
US-09-969-034-1012

Query Match 0.9%; Score 24; DB 11; Length 703;  
Best Local Similarity 100.0%; Pred. No. 0.45;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679  
DB 36 ATTGACAAAAA 13

RESULT 38  
US-10-437-963-69278/C  
Sequence 69278, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT FILING DATE: 2003-05-14  
CURRENT APPLICATION NUMBER: US/10/437,963  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 69278  
LENGTH: 724  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_69962C.1  
US-10-437-963-69278

Query Match 0.9%; Score 24; DB 17; Length 724;  
Best Local Similarity 100.0%; Pred. No. 0.45;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679  
DB 65 ATTGACAAAAA 42

RESULT 39  
US-10-444-795B-796  
Sequence 796, Application US/10444795B  
Publication No. US20040077544A1  
GENERAL INFORMATION:  
APPLICANT: Klinghoffer, Richard  
APPLICANT: Lewis, Stephen Patrick  
TITLE OF INVENTION: MODULATION OF BIOLOGICAL SIGNAL  
TITLE OF INVENTION: TRANSDUCTION BY RNA INTERFERENCE  
FILE REFERENCE: 200125,449  
CURRENT APPLICATION NUMBER: US/10/444,795B  
CURRENT FILING DATE: 2003-05-23  
NUMBER OF SEQ ID NOS: 842  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 796  
LENGTH: 844  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-444-795B-796

Query Match 0.9%; Score 24; DB 16; Length 844;  
Best Local Similarity 100.0%; Pred. No. 0.45;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679  
DB 821 ATTGACAAAAA 844

RESULT 40  
US-10-723-860-5727  
Sequence 5727, Application US/10723860  
Publication No. US20040253606A1  
GENERAL INFORMATION:  
APPLICANT: Aziz, Natsasha  
APPLICANT: Gineburg, Wendy M.  
APPLICANT: Zlotnik, Albert  
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
FILE REFERENCE: 05882,0193.NFUS01  
CURRENT APPLICATION NUMBER: US/10/723,860  
CURRENT FILING DATE: 2003-11-26  
PRIOR APPLICATION NUMBER: 60/429,739  
PRIOR FILING DATE: 2002-11-26  
NUMBER OF SEQ ID NOS: 8393  
SOFTWARE: Patentin version 3.2  
SEQ ID NO 5727  
LENGTH: 1486  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-723-860-5727

Query Match 0.9%; Score 24; DB 18; Length 1486;  
Best Local Similarity 100.0%; Pred. No. 0.45;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2656 ATTGACAAAAA 2679  
DB 884 ATTGACAAAAA 907

RESULT 41  
US-10-739-930-5455/C  
Sequence 5455, Application US/10739930  
Publication No. US20040216190A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-21(53377)B  
CURRENT APPLICATION NUMBER: US/10/739,930  
CURRENT FILING DATE: 2003-12-18  
NUMBER OF SEQ ID NOS: 11088  
SEQ ID NO 5455  
LENGTH: 1930  
TYPE: DNA  
ORGANISM: Triticum aestivum  
FEATURE:  
OTHER INFORMATION: Clone ID: TRIAR-23APR03-CLUSTER33891\_1  
US-10-739-930-5455

Query Match 0.9%; Score 24; DB 18; Length 1930;  
Best Local Similarity 100.0%; Pred. No. 0.45;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2654 AAATTGACAAAAA 2677  
DB 24 AAATTGACAAAAA 1

RESULT 42  
US-10-425-115-119860  
Sequence 119860, Application US/10425115

```
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5322)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 119860
LENGTH: 2192
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MT14577_40795C.1
US-10-425-115-119860

Query Match
Best Local Similarity 100.0%; Score 24; DB 18; Length 2192;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2656 ATTGACAAAAA 2679
2125 ATTGACAAAAA 2148

RESULT 43
US-09-764-869-2292
Sequence 2292, Application US/09764869
Patent No. US20020061521A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007
CURRENT APPLICATION NUMBER: US/09/764,869
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2292
LENGTH: 11869
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-869-2292

Query Match
Best Local Similarity 100.0%; Score 24; DB 9; Length 11869;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2260 CCCTCCCGAGAGCCCGGGGCGG 2283
1242 CCCTCCCGAGAGCCCGGGGCGG 1265

RESULT 44
US-10-091-504-2292
Sequence 2292, Application US/10091504
Publication No. US20030059908A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007C1
CURRENT APPLICATION NUMBER: US/10/091,504
CURRENT FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 2442
Prior application removed - See file wrapper or Palm
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2292
LENGTH: 11869
TYPE: DNA
```

```
ORGANISM: Homo sapiens
US-10-091-504-2292

Query Match
Best Local Similarity 100.0%; Score 24; DB 14; Length 11869;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2260 CCCTCCCGAGAGCCCGGGGCGG 2283
1242 CCCTCCCGAGAGCCCGGGGCGG 1265

RESULT 45
US-10-227-577-2292
Sequence 2292, Application US/10227577
Publication No. US20040005575A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007C2
CURRENT APPLICATION NUMBER: US/10/227,577
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/091,504
PRIOR FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: 09/764,869
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 2442
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2292
LENGTH: 11869
TYPE: DNA
ORGANISM: Homo sapiens
US-10-227-577-2292

Query Match
Best Local Similarity 100.0%; Score 24; DB 16; Length 11869;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2260 CCCTCCCGAGAGCCCGGGGCGG 2283
1242 CCCTCCCGAGAGCCCGGGGCGG 1265

RESULT 46
US-10-292-798-1095/C
Sequence 1095, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 08435/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
```

PRIOR APPLICATION NUMBER: 10/017,161  
PRIOR FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: JP 2001-246789  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 2070  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1095  
LENGTH: 358246  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
LOCATION: source  
FEATURE:  
LOCATION: (1)..(358246)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (201)..(326)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (605)..(1294)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (29165)..(29206)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (39051)..(39221)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (72539)..(72763)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (133049)..(133353)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (261460)..(261526)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (291173)..(291325)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (291513)..(291770)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (302538)..(302800)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (325058)..(325316)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (330441)..(330609)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (337750)..(337828)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (343822)..(344071)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (358008)..(358046)  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (547)..(646)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1659)..(1758)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (2818)..(2917)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base

LOCATION: (17179)..(17278)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (17497)..(17497)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (17499)..(17499)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (17518)..(17518)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (51382)..(51481)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (64360)..(64459)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (71021)..(71120)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (77979)..(78078)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (81210)..(81309)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (103441)..(103540)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (130963)..(130964)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (130976)..(130977)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (130979)..(130979)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (130984)..(130984)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (130987)..(130987)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (130997)..(130999)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (131002)..(131002)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (131019)..(131020)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (131022)..(131022)

```

; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131025)..(131025)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131031)..(131031)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131034)..(131034)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131048)..(131049)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131059)..(131060)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131063)..(131063)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131065)..(131065)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131070)..(131070)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131090)..(131091)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131127)..(131129)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base

```

```

Query Match      0.9%; Score 24; DB 15; Length 358246;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2656 ATTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      242855 ATTGACAAAAAAAAAAAAAAAAAAAA 242832

```

```

RESULT 47
US-10-242-535A-54104
; Sequence 54104, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54104
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-54104

```

```

Query Match      0.9%; Score 23; DB 16; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2657 TTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      130 TTGACAAAAAAAAAAAAAAAAAAAA 152

```

```

RESULT 48
US-10-085-783A-54104
; Sequence 54104, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54104
; LENGTH: 160
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-54104

```

```

Query Match      0.9%; Score 23; DB 16; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2657 TTGACAAAAAAAAAAAAAAAAAAAA 2679
DB      130 TTGACAAAAAAAAAAAAAAAAAAAA 152

```

```

RESULT 49
US-10-741-601-22073/C
; Sequence 22073, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; PRIOR FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22073
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-22073

```

```

Query Match      0.9%; Score 23; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2657 TTGACAAAAAAAAAAAAAAAAAAAA 2679  
Db 106 TTGACAAAAAAAAAAAAAAAAAAAA 84

## RESULT 50

US-09-867-701-9978  
; Sequence 9978, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Agiate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9978  
; LENGTH: 222  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-9978

Query Match 0.9%; Score 23; DB 9; Length 222;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2657 TTGACAAAAAAAAAAAAAAAAAAAA 2679  
Db 177 TTGACAAAAAAAAAAAAAAAAAAAA 199

Search completed: January 15, 2005, 12:52:34  
Job time : 1458 secs